Summer 7-12-2012

Two-site DNA Minor Groove Binding Compounds

Shelby Diane Sheldon Deuser
Georgia State University, ssheldon4@student.gsu.edu

TWO-SITE DNA MINOR GROOVE BINDING COMPOUNDS

By

SHELBY DIANE SHELDON DEUSER

Under the Direction of Dr. W. David Wilson

ABSTRACT

DNA minor groove binding compounds have had limited therapeutic uses, in part due to problems with sequence specificity. A two-site model has been developed to enhance specificity, in which compounds bind to two short AT sites separated by one or two GC base pairs. Using thermal melting, heterocyclic dications with this capability were tested with various oligonucleotides for binding affinity and specificity. Compounds of interest were further probed using circular dichroism, mass spectrometry, biosensor-SPR, and molecular modeling. Several compounds were found to “jump” a GC base pair, binding to AT sites in the minor groove of DNA with a two-site recognition mode. One compound was also found to recognize a single intervening GC base pair. Compounds with terminal, non-polar amidine extensions were found to have increased DNA binding compared to analogs with terminal amidines. This unique, two-site DNA recognition mode offers novel design principles to recognize entirely new DNA motifs.

INDEX WORDS: DNA, Minor groove binders, Heterocyclic diamidines, Trypanosomiasis, Two AT site complexes, Sequence recognition, Thermal melting, Surface plasmon resonance, Circular dichroism, Mass spectrometry

TWO-SITE DNA MINOR GROOVE BINDING COMPOUNDS

by

SHELBY DIANE SHELDON DEUSER

A Thesis Submitted in Partial Requirement for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2012

TWO-SITE MINOR GROOVE BINDING COMPOUNDS

by

SHELBY DIANE SHELDON DEUSER

Committee Chair: Dr. W. David Wilson

Committee: Dr. David Boykin

Dr. Markus Germann

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2012

DEDICATION

Dedicated to my husband Karl, whose encouragement, love, and support have been unwavering.

ACKNOWLEDGEMENTS

This work would not be possible without the patient guidance of Dr. W. David Wilson, who has been an invaluable mentor and leader. Special thanks to Dr. Rupesh Nanjunda for answering strange questions, discussing unexpected results, and teaching me the technique of surface plasmon resonance. The molecular modeling was made possible with the help of Dr. Michael Rettig. Every member in the Wilson lab group has supported this project in some way, and I am honored to have worked in the Wilson lab.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Nucleic Acids as Drug Targets

The human body is based upon the language of molecules, and through their interactions states of wellness and disease are created. The effects, and often the root causes of disease can be traced back to aberrant transcriptional pathways. The biological system of transcription is complex and dynamic, consisting of proteins, nucleic acids, co-factors, and other compounds whose interactions maintain life. By targeting nucleic acids in a sequence-specific manner with small, cell-permeable molecules, cellular regulation and genomic reprogramming may become medicines of the future.[1]

With the advent of the human genome project, nucleic acids have become an increasingly popular target, both for basic research and for novel therapeutics. Current and early anti-cancer drugs target DNA, including actinomycin D and cisplatin.[2] The treatment for tuberculosis, streptomycin, specifically targets ribosomal RNA in the pathogenic microorganism.[2b] Aromatic diamidines such as pentamidine, currently used as treatment, target the minor groove of kinetoplast DNA (kDNA) in the trypanosome parasite responsible for African sleeping sickness. These circular kDNAs are interlocked similarly to chain mail, and are found inside the protozoan's mitochondria. Since there are no equivalent nucleic acid structures in humans, these sites are attractive targets for antiparasitic drug development.[2b, 3] The kinetoplast sites are rich in AT base pairs, where classical heterocyclic dications prefer to bind.[3a, 3c, 4] However, human DNA also

contains sites rich in AT base pairs; thus greater sequence specificity must be developed to increase the therapeutic index of this class.

1.2 Two-site Approach

The AT sites found in kinetoplast minicircles of mitochondrial DNA are often separated by one or more GC base pairs (BPs); a common motif is 3-4 AT BPs separated by 1-2 GC BPs.[3a] If this sequence-specific feature could be selectively targeted, minor groove binding agents could be improved as anti-parasitic agents with fewer side effects. The minor groove of an AT BP (Figure 1) has two hydrogen bond acceptor sites, the N3 of adenine and the C2 carbonyl of thymine; and is easily recognizable by heterocyclic groups with hydrogen bond donors.[4-5] The minor groove of a GC BP has one hydrogen bond donor group, the C2 amine group of guanine, and two acceptors, the N3 of guanine and the C2 carbonyl of cytosine. However, the protruding exocyclic amine group of guanine presents a steric hindrance to deep minor groove binding, and recognition is challenging.

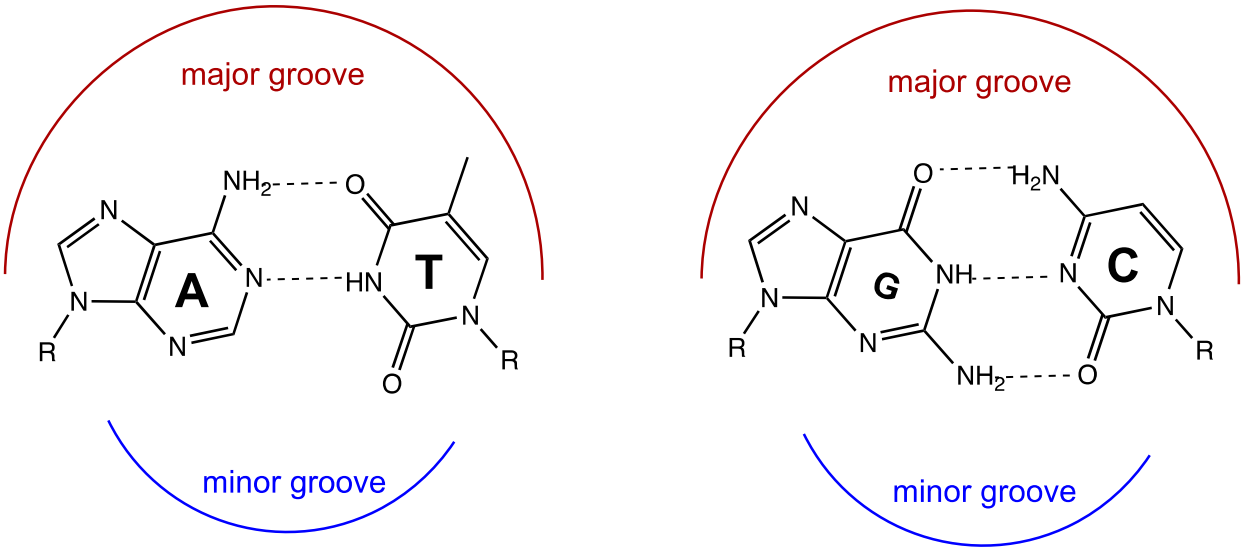


Figure 1. Structure of DNA base pairs.[4] Adenine-thymine (AT) and guanine-cytosine (GC) are illustrated with major and minor grooves indicated. Note the exocyclic amine of guanine protruding into the minor groove.

AT-rich sequences, especially poly-dA oligonucleotides, have other unique characteristics. The minor grove of such sequences is compressed, resulting in increased negative electrostatic potential.[4, 6] Introduction of alternating A-T "steps" lessens this effect. If intervening GC BPs are introduced, the minor groove widens, and the resulting electrostatic potential is relatively decreased in magnitude. Because of these minor groove characteristics, preferential binding to GC BPs is both sterically and electrostatically challenging. By linking two AT recognizing moieties (Figure 2), compounds that "jump" the GC BPs can be designed. A linker capable of recognizing the GC BPs is a future goal.

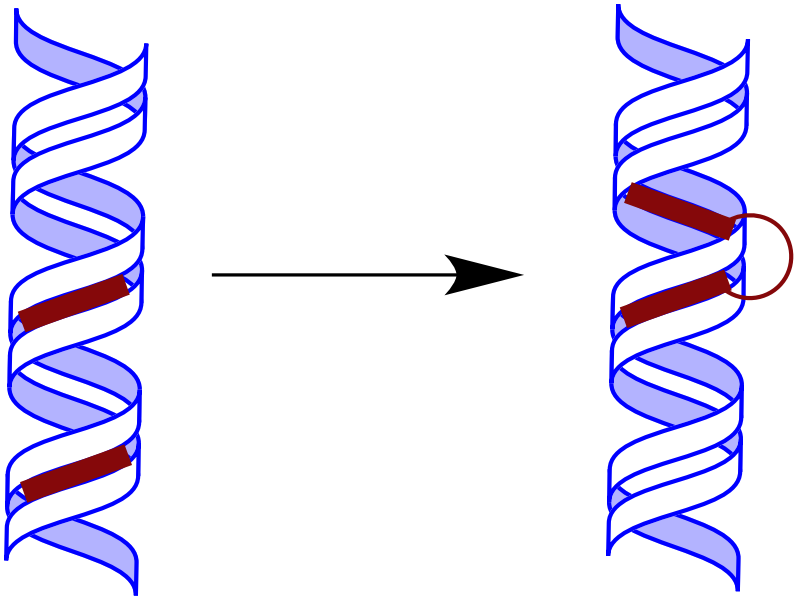


Figure 2. Two-site concept. Initially, AT-recognizing compounds bind 3-4 AT base pairs each. The two are joined together with a linker, and can span 1-2 intervening GC base pairs for a longer cognate binding site.

A two-site approach has been used with a variety of DNA binding systems. Pyrrolo[2,1-*c*][1,4]benzodiazepines (PDBs) such as tomamycin and anthramycin are antitumor antibiotic agents currently used clinically.[7] These compounds bind to the minor groove, covalently bonding to the exocyclic C_2-NH_2 of a guanine base. Two such PDBs have been joined via a diether linkage (*i.e.* -O-$(CH_2)_n$-O-), and various homologs tested for linker length optimization. The homolog studies show that three and five methylene groups correlate with guanine separation of two and three BPs respectively, with the two covalent bonds forming on opposite strands. Similarly, to serve as a probe for higher-ordered DNA structures, two Hoechst 33258 moieties have been joined with an oligo-ethylene glycol linker.[8] Hoechst 33258 is known to strongly bind to the minor groove of DNA at an AT rich site (A_3T_3). The bidentate compounds bind to the minor groove of DNA at two AT sites, showing cooperative binding as well as site discrimination.

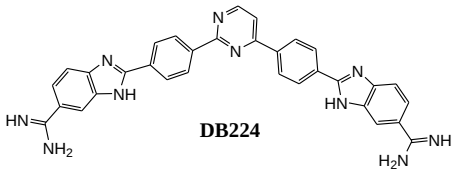
Other examples include polyamides, minor groove binders based on the natural products netropsin and distamycin A, which can be programmed for binding in a sequence specific manner.[1a, 9] Hairpin polyamides have been connected together head-to-head via alkyl linkers (*i.e.* -CO(CH_2)$_n$CO-). The resulting compounds bind one hairpin moiety to each of two adjacent binding sites, spanning 10-11 BP.[9c] Linker optimization studies showed that two methylene groups are preferred for jumping two AT BPs. Polyamides have also been developed that bind as overlapped dimers.[10] These compounds break the cognate sequence into two halves, binding to AT rich opposite sides of the binding site and stacking in the mixed BP center. Known as combilexins, compounds have also been designed that combine a polyamide-type minor groove binder with an intercalating agent, resulting in anticancer molecules that bind with two different modes to two different sites.[11]

Successful implementation of a two-site model would result in compounds that are thermodynamically driven towards complex formation with longer binding sites. Binding with a single site would leave half of the ligand unbound, resulting in a high energy, unstable complex. Such selectivity promises enhanced recognition, which is necessary for genome-regulating compounds as well as for a new generation of anti-kDNA therapeutics.

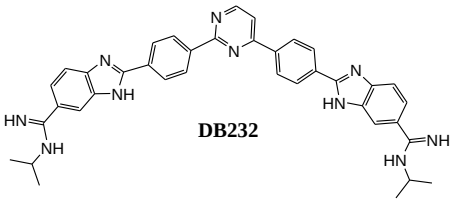
1.3 Purpose of This Study

Minor groove DNA binding compounds hold great potential for new drugs, not only as anti-parasitics, but also as human genome regulators that can target other diseases. For this to become a reality a greater understanding of compound structure vs. selectivity must be developed. In order to keep the synthetic price of potential therapeutics relatively low, this study focuses on symmetrical molecules. Although sequence-specific polyamides can be designed and synthesized, they are currently expensive and impractical as new drug candidates. Other minor groove binding heterocyclic groups such as benzimidazole and hydroxybenzimidazole have been developed that show promising selectivity;[1a, 9b] however, no true substitution for polyamide sequence recognition has yet been discovered. This study screens a number of heterocyclic dications (Figure 3) against a variety of DNA sequences (Figure 4). From the resulting matrix of data, structure-activity relationships can be derived that further this goal. The sequence $(AAG)_3$ is a specific target for binding. This sequence is repeated many times in the genetic disease Friedreich's Ataxia, causing irregular DNA hairpin formations. The body's repair mechanisms are ineffective at removing these repeats, resulting in "sticky" DNA that forms a triplex structure.[12] Stabilizing the duplex form of this DNA by minor groove binding molecules is a novel treatment approach. Unlike many minor groove binders, heterocyclic dications are readily cell-permeable; increased selectivity will open the door for new drug development.

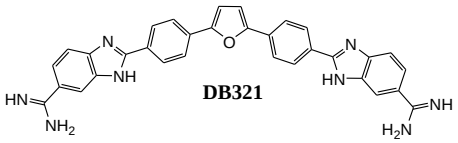
Group I



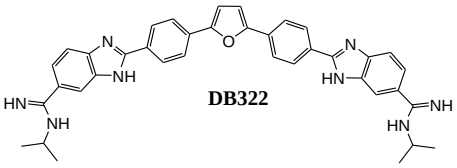
DB224



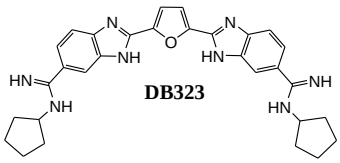
DB232



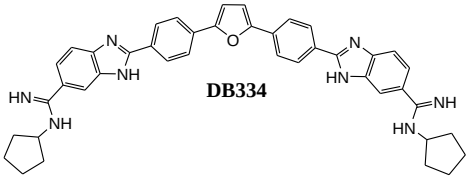
DB321



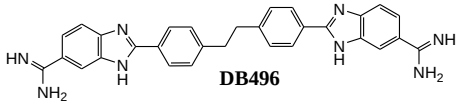
DB322



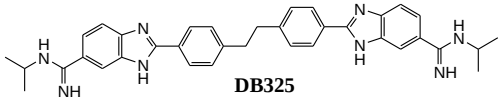
DB323



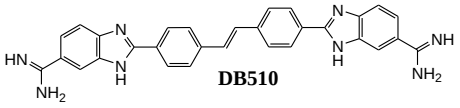
DB334



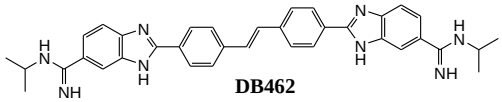
DB496



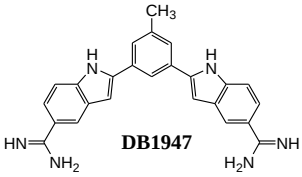
DB325



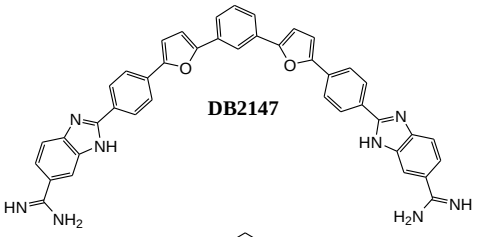
DB510



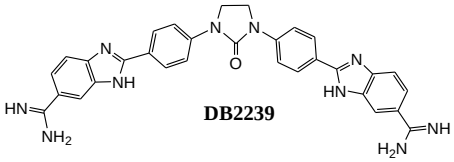
DB462



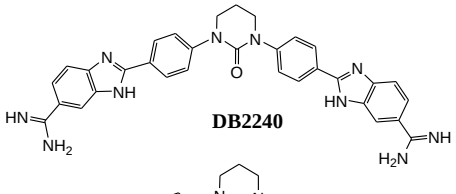
DB1947



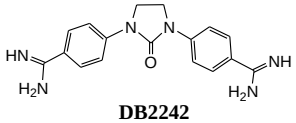
DB2147



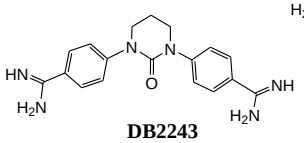
DB2239



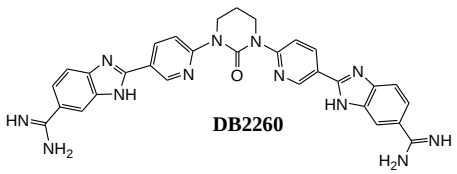
DB2240



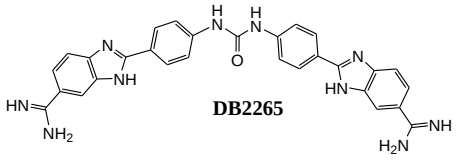
DB2242



DB2243

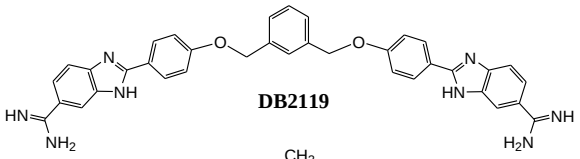


DB2260

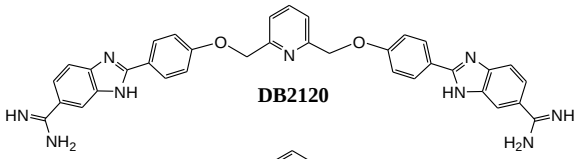


DB2265

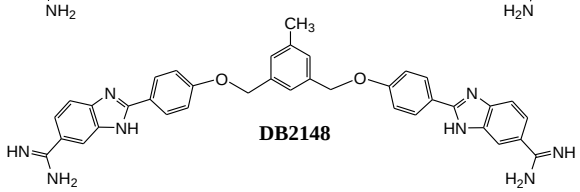
Group II



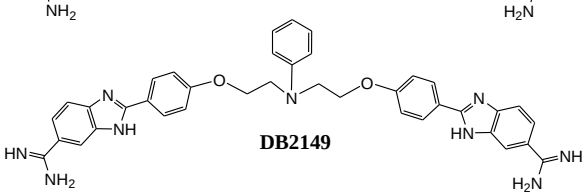
DB2119



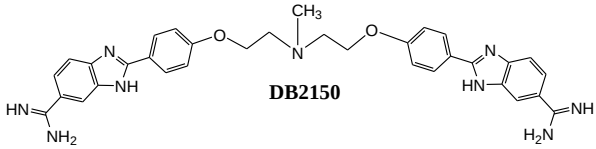
DB2120



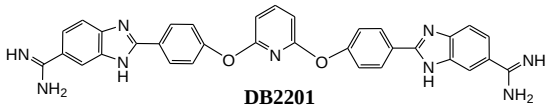
DB2148



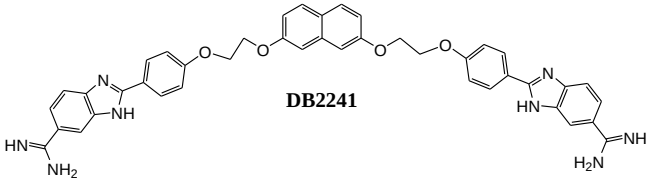
DB2149



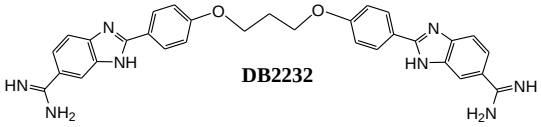
DB2150



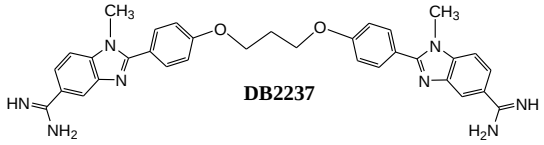
DB2201



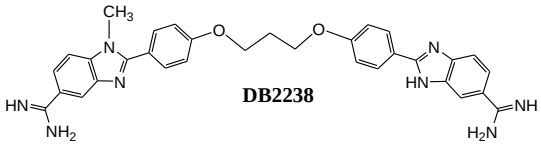
DB2241



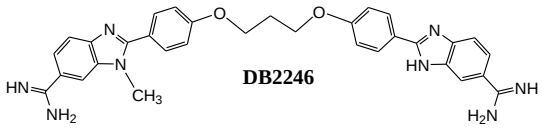
DB2232



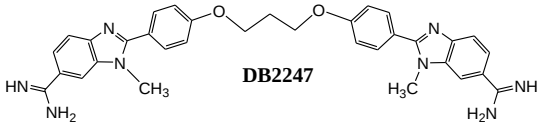
DB2237



DB2238

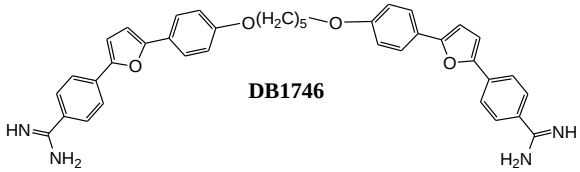


DB2246

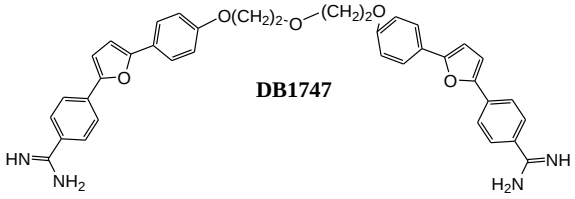


DB2247

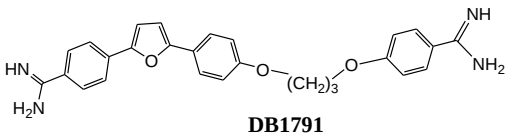
Group III



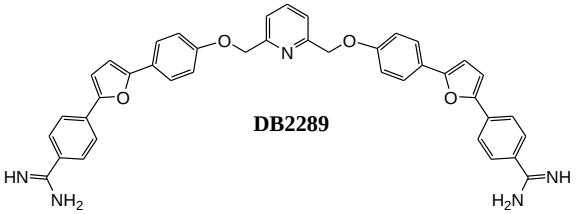
DB1746



DB1747



DB1791



DB2289

Group IV

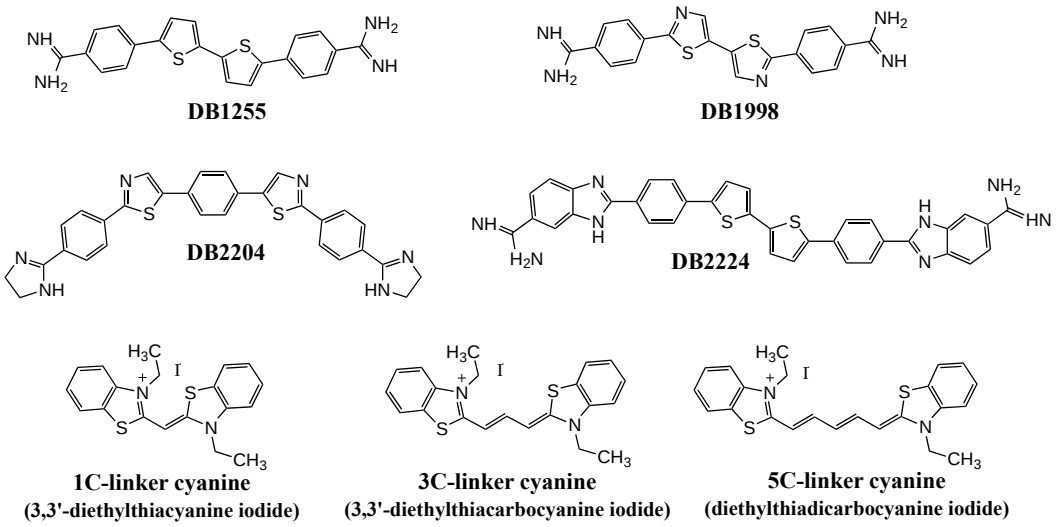


Figure 3. Heterocyclic diamidine and cyanine dye structures. Group I compounds consist of conformationally limited, conjugated ring systems with predominantly phenyl-benzimidazole-amidine terminals; Group II compounds consist of similar systems with a more flexible, central region connected via diether linkage; Group III compound consist of compounds with a longer, flexible linker region and phenyl-furan-phenyl terminals; and Group IV compounds consist of conformationally limited, conjugated ring systems with sulfur introduced.

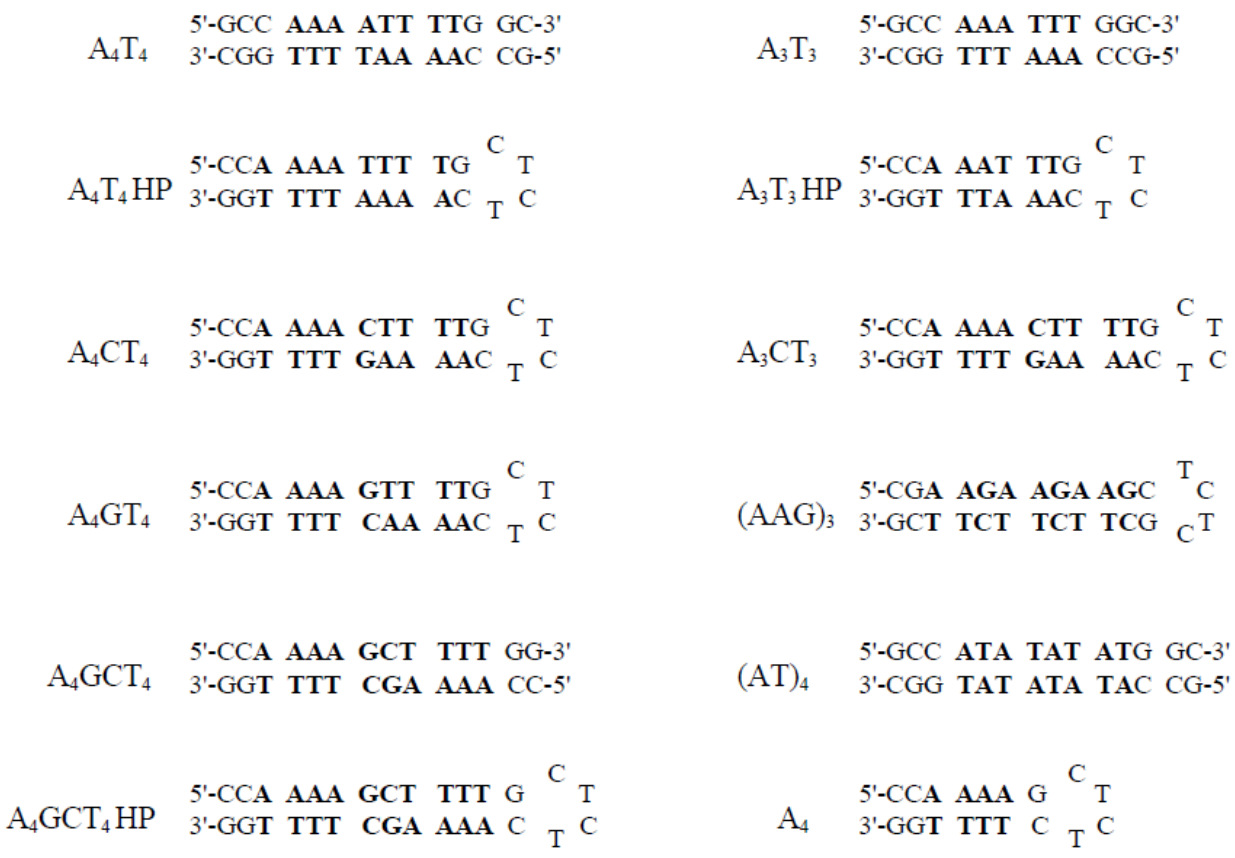


Figure 4. DNA seuqences tested. Shorthand is listed to the left of each oligonucleotide, and will be used to refer to the sequences throught the text. Hairpin and self-complementary DNA sequences were selected with a variety of binding sites and similar flanking sequences. A_4T_4 is the two-site sequence with no intervening GC BPs, and A_3T_3 the shortened version, selected as a test for binding site length. A_4 is the single-site control. A variety of two-site sequences with one or two intervening GC BPs were selected (A_4GT_4, A_4CT_4, A_4GCT_4, and A_3CT_3), as well as the Friedreich's Ataxia sequence, $(AAG)_3$, which can be thought of as a shortened two-site or three-site sequence. The sequence $(AT)_4$ was selected to compare an A-T step sequence to the A-tract sequences.

2 EXPERIMENTAL

2.1 Buffer and Sample Preparation

2.1.1 Cacodylic Acid Buffer

Cacodylic acid (CCL), disodium ethylenediamine tetraacetic acid (EDTA), sodium hydroxide (1 N), and sodium chloride were obtained from Fisher Scientific (Fairlawn, NJ). Nanopure water was obtained using an ELGA Purelab Classic water purification system. Cacodylic acid (10 mM), NaCl (100 mM), and EDTA (1 mM) were dissolved into nanopure water by stirring over low heat. An Accumet pH Meter 910 (Fisher Scientific) was used to monitor pH, and the buffer was titrated using NaOH until the pH reached 6.2. The solution was transferred into a Kimax volumetric flask and made up to the mark with additional nanopure water.

2.1.2 2[N-Morpholino]ethanesulfonic Acid Buffer

Surfactant P20 was obtained from GE Healthcare (Uppsala, Sweden), 2[*N*-morpholino]ethanesulfonic acid (MES) was obtained from Sigma Chemical Company (St. Louis, MO), and sodium hydroxide was obtained from Fisher Scientific (Fairlawn, NJ). MES (10 mM), NaCl (100 mM) and EDTA (0.1 mM) were dissolved into nanopure water by stirring over low heat. The solution was titrated with NaOH to a pH of 6.2, and the total volume made up in a volumetric flask as described above. Surfactant P20 (0.005% or 0.05%, v/v) was added, and the solution filtered and degassed using a

vacuum filtration apparatus with 0.22 µm nitrocellulose filter paper (Millipore, Billerica, MA).

2.1.3 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic Acid Buffer

Hydrochloric acid (1 N) was obtained from Fisher Scientific (Fairlawn, NJ), and 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) was obtained from Sigma Chemical Company (St. Louis, MO). HEPES (10 mM), NaCl (100 mM) and EDTA (0.1 mM) were dissolved into nanopure water by stirring. The solution was titrated with hydrochloric acid to a pH of 7.4, and the total volume made up in a Kimax volumetric flask. Surfactant P20 (0.05% v/v) was added, and the solution filtered and degassed as described above.

2.1.4 Compound Preparation

All DB compounds were synthesized in Professor David Boykin's laboratory (GSU, Atlanta, GA). Cyanine dyes (3,3'-diethylthiadicarbocyanine iodine, 3,3'-diethylthiacyanine iodide, and 3,3'-diethylthiacarbocyanine iodide) were obtained from Aldrich Chemical Company (Milwaukee, WI). Stock solutions of compounds (0.3 - 1 mM) were prepared in nanopure water with 0-30% DMSO added if necessary for dissolution. Extinction coefficients for selected compounds were obtained using a Cary 300 Bio UV-vis spectrophotometer (Varian) run from 800-200 nm at 60 nm/min. A plot of concentration vs. peak absorbance was created using Kaleidagraph software, and the slope taken to be the extinction coefficient of the compound as evident from the Beer-Lambert law.

2.1.5 Oligonucleotide Preparation

Oligonuleotides were obtained from Integrated DNA Technologies (Coralville, IA). Solutions were prepared by adding nanopure water to the oligonucleotide vial for ~1 mM concentration. These concentrations were checked by using the simple reads function of a Cary 300 Bio UV-vis spectrophotometer (Varian) at 260 nm. The absorbance values of five titrations of DNA stock solution were averaged and used to determine concentration based on the Beer-Lambert Law.

2.2 Thermal Melting

Binding of compounds to oligonucleotides was screened by thermal melting using a Cary 300 Bio UV-vis spectrophotometer (Varian). Samples of DNA and compound in buffer (1000 µL) were prepared in 1 cm quartz cuvettes at various ratios (*e.g.* 0:0 (buffer blank), 0:1 (DNA), 1:1, 2:1 and 3:1), and were annealed prior to being tested. For bulk screening purposes, different compounds at a 2:1 ratio were tested at the same time. A concentration of 3 µM DNA was used unless significant aggregation occurred, in which case a concentration of 1 µM DNA was used. The machine was set to test a single ramp at 260 nm and 0.5° C/min, beginning well below the native melting temperature and ending well above it or at 95.00° C. The absorbance of the buffer was subtracted, and a graph of normalized absorbance vs. temperature was created using Kaleidagraph software. The ΔTm values were calculated using a combination of the derivative function and estimation from the normalized graphs.[13]

2.3 Circular Dichroism

Circular dichroism studies were performed using a Jasco J-810 instrument. Four spectra were averaged from 500 to 220 nm, using a 1 cm quartz cuvette, a scan speed of 50 nm/min, and a response time of 1 sec. A cacodylic acid buffer spectrum was taken first, followed by DNA in the same buffer (3 μM). Compound was then titrated into the solution, taking a spectrum at each new concentration until saturation was reached. Buffer subtracted graphs were created using the Kaleidagraph software.

2.4 Mass Spectrometry

Hairpin oligonucleotide solutions were desalted three times using 150 mM ammonium acetate (Sigma-Aldrich, St. Lious, MO) buffer and a 1000 Da cut-off membrane (Spectrum Laboratories Inc., Rancho Dominguez, CA). A concentration of 5 μM of each oligonucleotide was run together with an appropriate amount of compound to yield the desired ratio, reported as [compound]:[each oligonucleotide]. Compounds were prepared in ammonium acetate buffer (150 mM) with or without 5% (v/v) methanol (Fisher Scientific, Fairlawn, NJ). Electrospray ionization mass spectrometry on a negative mode was performed using a Waters Micromass Q-TOF micro (ESI-Q-TOF, Milford, MA) instrument. ‘Soft’ ionization conditions were used, with a capillary voltage of 2.5 kV, sample cone voltage of 30 V, source block temperature of 70° C, desolvation temperature of 100° C, nitrogen as the nebulizing and drying gas, and direct injection at 5 μL/min. Spectral information was recorded in the mass/charge region 300-3000, and MassLynx 4.1 software used for data analysis.

2.5 Surface Plasmon Resonance

Surface plasmon resonance (SPR) was conducted using a four-channel Biacore 2000 or Bia T200 optical biosensor system (Biacore, GE Helthcare Inc.). Hairpin oligonucleotides biotin labeled at the 5' end were immobilized on to streptavidin-coated sensor chips (Biacore SA) as described in the literature.[1b, 2b, 14] One flow cell was left blank as a reference, and the three remaining cells immobilized with ~350 RU DNA. Filtered, degassed MES or HEPES buffer was used for all SPR experiments. Serially diluted compound samples were prepared and tested, and the resulting sensorgrams fit using steady state models.[2b, 14a] Data processing was performed using the BIAevaluation software (Biacore, GE Heathcare Inc.) and Kaleidagraph software.

2.6 Molecular Modeling

Compounds were modeled using Spartan '04 running on a Windows XP platform. Molecules were drawn and the molecular geometry minimized with molecular mechanics based on the MMFF force field after the addition of each group. Next, an equilibrium conformer calculation was performed using the MMFF force field. This output was further refined by calculating the equilibrium geometry using a Hartree-Fock model with a 6-31G* basis set. A high-resolution electrostatic potential map was generated by mapping the electrostatic potential onto an electron density surface.[15]

3 RESULTS

3.1 Extinction Coefficients

UV-vis absorption spectra were measured to determine if the compounds studied resulted in linear Beer-Lambert plots, as well as to determine extinction coefficient values. An example of spectral series and extinction coefficient are shown in Figure 5 for DB321. Despite their hydrophobicity, compounds typically resulted in linear Beer-Lambert plots throughout the concentration range tested (1-20 µM), with extinction coefficients in the range of 4 x 10^4 - 7 x 10^4 L/(M*cm), and peak absorbances between 270 and 400 nm.

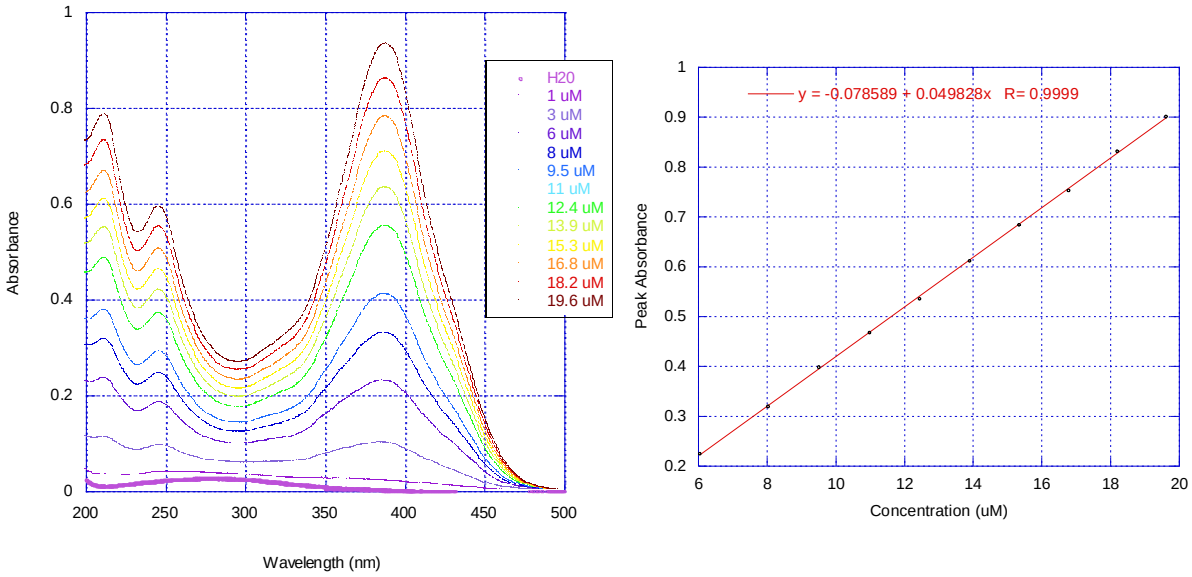


Figure 5: Extinction coefficient graphs for DB321. For DB321, a peak absorbance value at 381 nm can be found from the plot of absorbance vs. wavelength. As with the other compounds, the extinction coefficient was found from the slope of the best-fit line of a plot of peak absorbance vs. concentration.

Table 1: Extinction coefficients and peak absorbances of selected compounds. The extinction coefficient values were calculated using the Beer-Lambert Law, measuring the absorbance of various ratios of compound at the wavelength with maximum absorbance.

Compound	Extinction Coefficient (L/(M*cm))	Peak Absorbance (nm)
DB321	5.0×10^4	381
DB462	6.5×10^4	365
DB510	5.2×10^4	366
DB1947	6.8×10^4	277
DB2148	4.4×10^4	313

3.2 Thermal Melting

Thermal melting experiments were performed as a robust screening method for relative binding affinity. The method is illustrated in Figure 6, with different compounds and different DNA sequences.

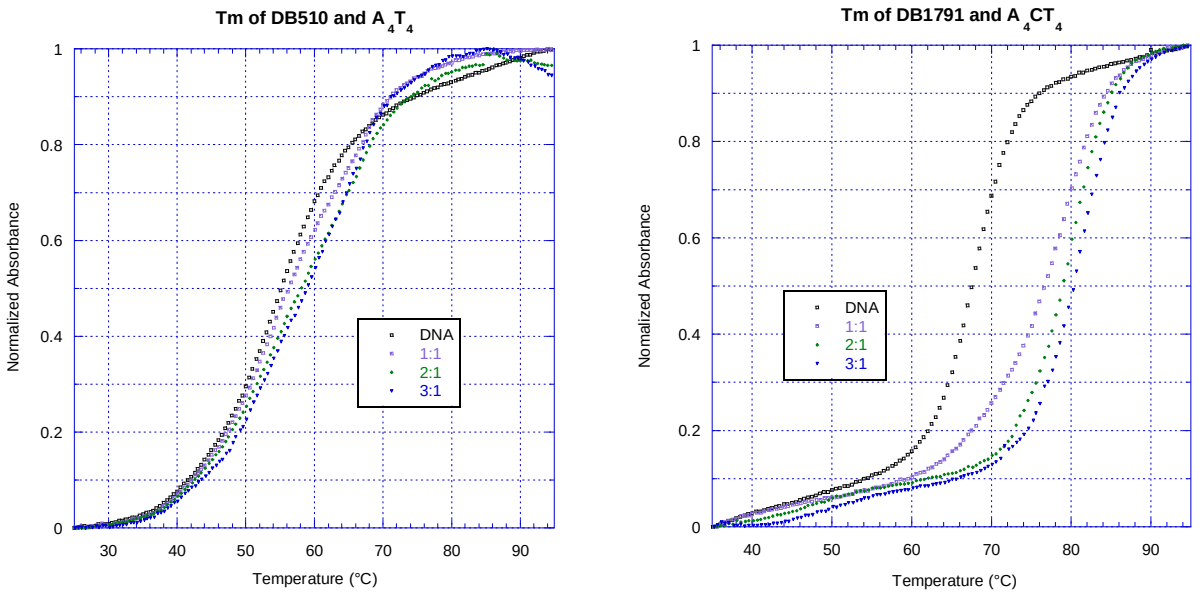


Figure 6. Selected thermal melting graphs (see Appendix for additional graphs). The change in thermal melting temperature was found by melting samples of DNA with compound at various ratios, and examining the change in the graph of normalized absorbance vs. temperature. Samples were typically prepared with 3 μM DNA, in either

CCL or MES buffer at pH 6.2. Typically a ratio of 2:1, [compound]:[DNA], was used to compare relative binding affinities. For compounds with abnormal graphs the derivative function occasionally failed to produce an accurate Tm value. In these cases, the ΔTm value was estimated from the graph.[13]

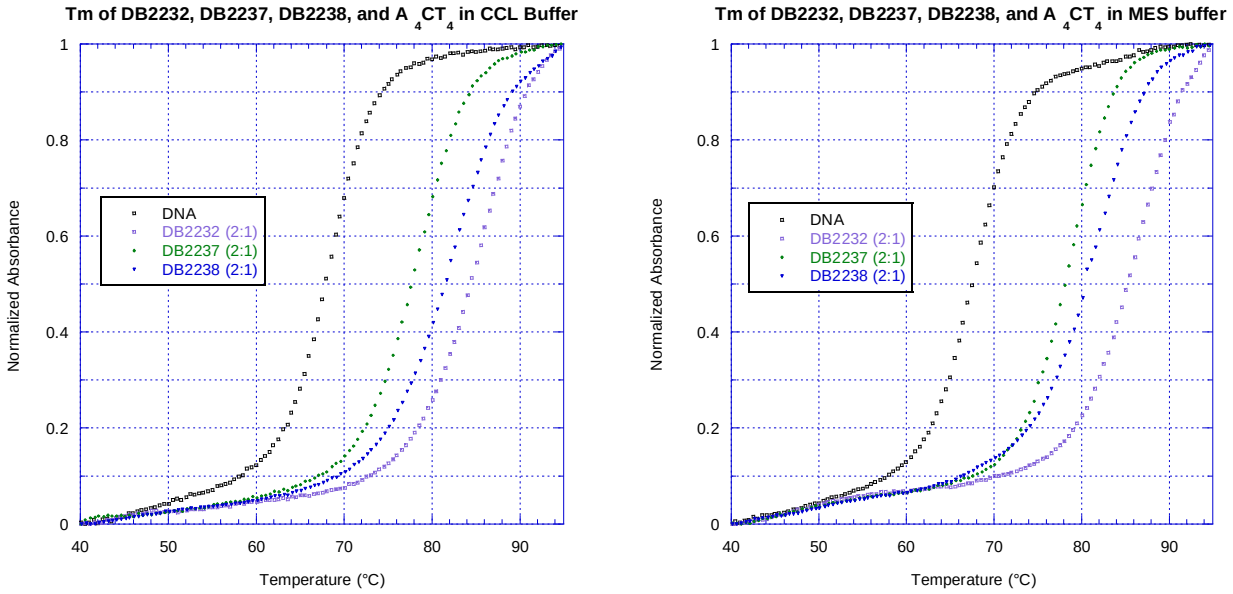


Figure 7. Comparison of CCL and MES buffers using thermal melting. Samples were tested with 3 μM DNA in either CCL or MES buffer at pH 6.2. Compounds were assumed to bind as either monomers or dimers, and so a 2:1 ratio was chosen for comparison. The ΔTm results for A_4CT_4 with DB2232, DB2237, and DB2238 were the same in CCL and MES buffers, indicating that these buffers can be used equivalently in thermal melting experiments.

Table 2: ΔTm results of various compounds and oligonucleotides. Compounds were tested with 3 μM or 1 μM DNA in either CCL or MES buffer (100 mM NaCl) at pH 6.2. ΔTm was calculated in degrees centigrade from (Tm with compound) - (Tm without compound) at a 2:1 ratio of compound to duplex DNA, and is assumed to have an error of +/- 1° C. "*" indicates trials run with hairpin sequences; "**" indicates that a 3:1 ratio was used for this compound; "---" indicates that no trial has been done.

	DNA	A_4T_4	A_4GT_4	A_4CT_4	A_3CT_3	A_3T_3	A_4GCT_4	$(AAG)_3$	$(AT)_4$	A_4
	Native Tm (° C)	55, 65*	67	66	66	54, 63*	53	73*	58	69
Group I	**DB224**	21	18	17	16	>17*	0	6	5	3
	DB232	26	>22	>22	>22	24	14	15	7	3
	DB321	23	23	22	23	5	0	16	5	2
	DB322	28	>25	>23	>22	29	20	8	8	8
	DB323	14	10	10	8	16*	9	6	-2	11
	DB334	27	>23	>21	>21	>25*	6	7	12	8
	DB325	>22	12	12	11	19*	10	5	1	7
	DB496	4	2	2	2	0*	2	2	-2	2
	DB462	13	10	10	7	14*	2	2	3	4
	DB510	4	2	1	2	3*	0	0	0	1
	DB2147	14	---	0	---	---	0	1	---	---
	DB1947**	16	13	14	15	18*	14	9	---	12
	DB2239	2*	3	2	1	2	1	1	-2	1
	DB2240	13*	10	9	12	11	7	4	1	3
	DB2242	2*	1	---	---	---	2	---	---	---
	DB2243	0*	1	---	---	---	0	---	---	---
	DB2260	15*	13	13	12	12*	11	5	0	7
	DB2265	12*	2	3	1	2	0	1	0	1
Group II	**DB2119**	14*	12	11	2	2	2	1	3	1
	DB2120	15*	>21	>21	>19	3	9	4	4	2
	DB2148	17	17	17	8	2*	1	2	2	0
	DB2149	11	5	6	3	2*	3	1	---	2
	DB2150	15*	14	13	13	14*	9	9	0	7
	DB2201	5*	4	---	---	---	1	2	---	---
	DB2241	2*	1	---	---	---	0	---	---	---
	DB2232	>24*	16	17	16	18	13	9	9	9
	DB2237	17*	9	10	9	13	8	4	-1	5
	DB2238	20*	13	13	11	15	10	6	3	7
	DB2246	20*	14	14	13	17	11	6	3	7
	DB2247	14*	8*	9	6	13*	7	3	-1	4

Group III	**DB1746**	2	0	0	---	---	2	---	---	---
	DB1747	3	---	3	---	---	0	---	---	---
	DB1791	15	13	14	13	14	9	5	0	4
	DB2189	2*	1	---	---	---	2	---	---	---
Group IV	**DB1998**	18	13	---	13	19	11	9	2	9
	DB1255	17	14	---	14	16	6	5	2	6
	DB2204	4	4	---	---	4*	4	---	3	---
	DB2224	0	0	---	---	---	-2	1	-2	---
	1C-Linker Cyanine	2	0	---	---	---	0	---	---	---
	3C-Linker Cyanine	2	1	---	---	---	1	---	---	---
	5C-Linker Cyanine	1	1	---	---	---	0	---	---	---

Many of the larger compounds showed preferences for sequences with two AT-rich sites, suggesting a two-site binding mode. Group I compounds elevated the melting temperature of such sequences the most, with ΔTm values over 22° C. All of the larger compounds that showed binding predominantly preferred sequences having two A-tracts, and had trouble accommodating two intervening GC base pairs. DB1947, a smaller compound, bound to the two-site sequences as well as to the others, showing little to no selectivity. DB2147 thermally stabilized the A_4T_4 sequence, but did not do the same to any of the other oligonucleotides. When compared to the A_4T_4 sequence, the introduction of alternating AT base pairs, $(AT)_4$, effectively decreased the binding of all compounds. Several compounds were ineffective as minor groove binders, including DB2242, DB2243, DB2201, the symmetrical compounds in Group III, the larger sulfur-containing compounds, and the selected cyanine dyes. However, many more

compounds were found to be strong binders by thermal melting, and show promising results.

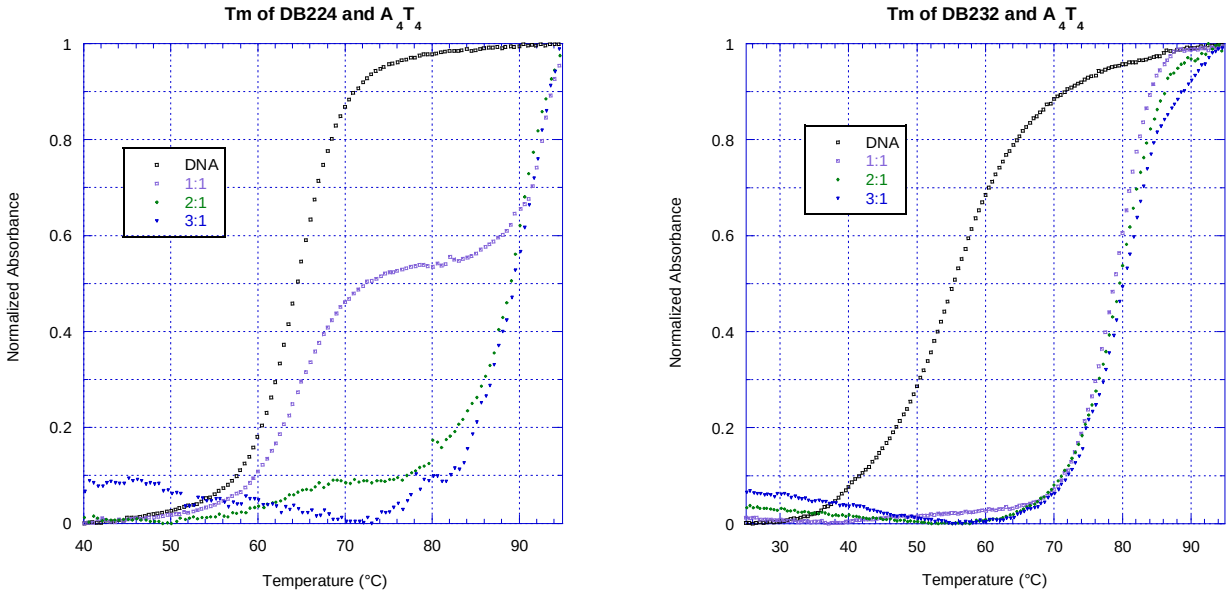


Figure 8. ΔTm graphs for DB224 and DB232 with A_4T_4. Compounds were tested at 0:1, 1:1, 2:1, and 3:1 ratios with 3 µM DNA in CCL or MES buffer (100 mM NaCl) at pH 6.2. DB232, the close analog of DB224, resulted in a more uniform thermal melting curve than did the parent compound. The 1:1 ratio of DB224 and A_4T_4 showed free DNA left in solution, while the 1:1 ratio of DB232 and A_4T_4 did not.

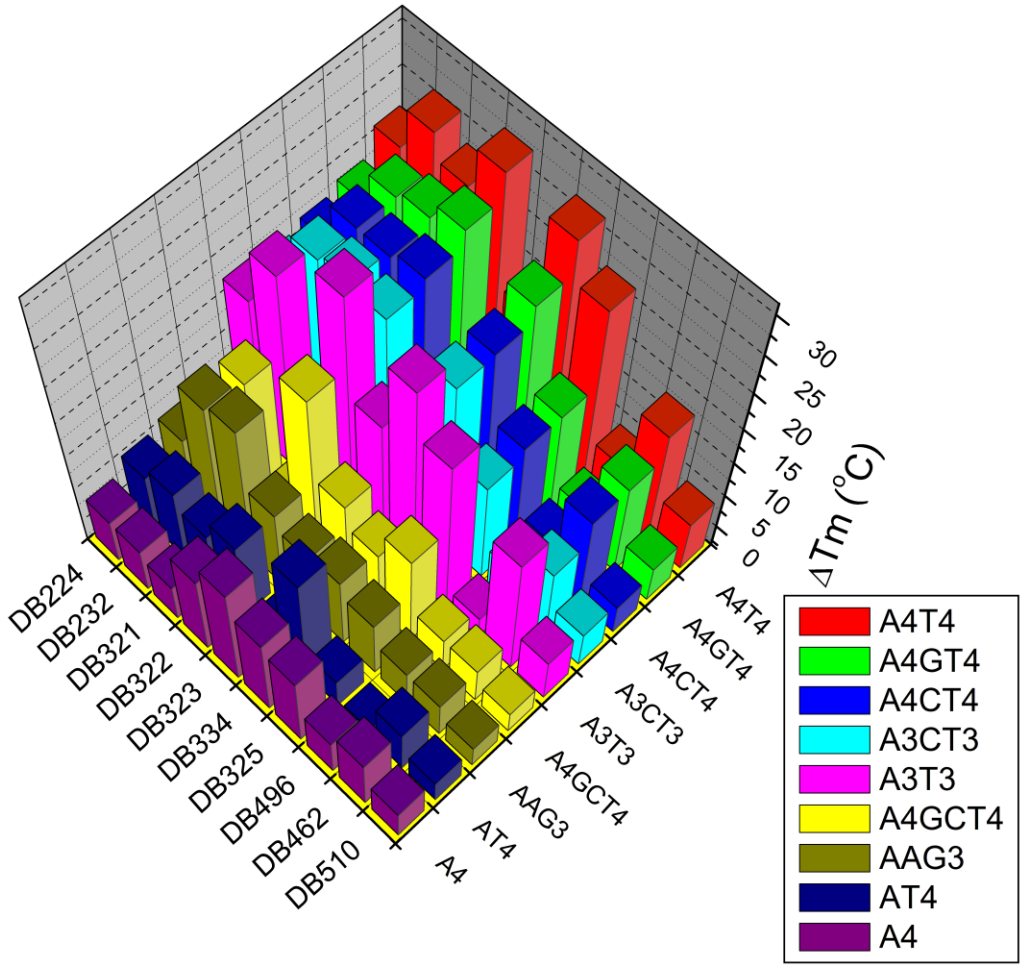


Figure 9. ΔTm results for DB224, DB321, DB496, DB510, and analogs. Compounds were tested at a 2:1 ratio with 3 µM DNA in CCL or MES buffer (100 mM NaCl) at pH 6.2. DB224, DB321, DB496, and DB510 bound to nearly all sequences more weakly than their analogs, DB232, DB322, DB334, and DB496, which are identical except for the inclusion of terminal non-polar groups. DB323, containing no phenyl groups, bound more weakly than DB321 to most sequences, exceptions being A_4, A_3T_3, and A_4GCT_4. Compounds predominantly preferred the A_4T_4 sequence, with most binding with similar affinity to the A_4GT_4, A_4CT_4, A_3CT_3, and A_3T_3 sequences.

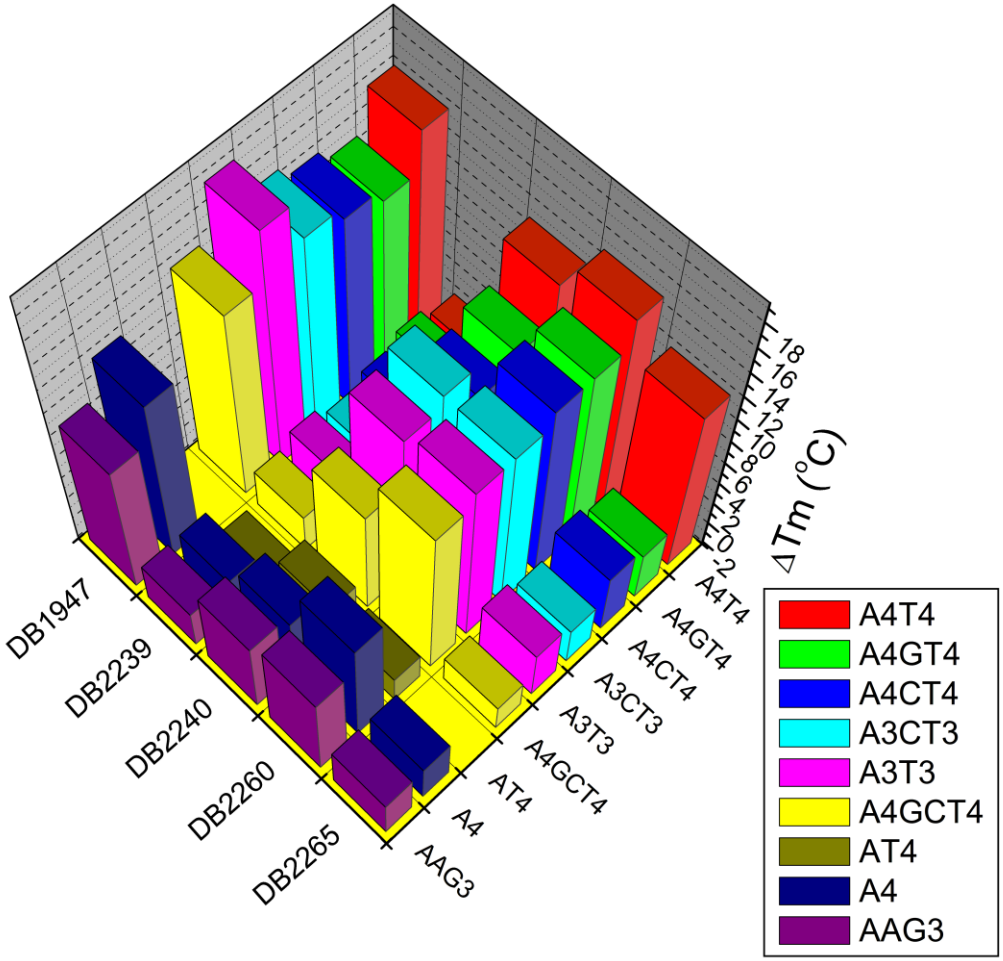


Figure 10. ΔTm results for remaining Group I compounds. Compounds were tested at a 2:1 ratio, with the exception of DB1947, which was tested at a 3:1 ratio. A concentration of 3 µM DNA was used in CCL or MES buffer (100 mM NaCl) at pH 6.2. Again, compounds predominantly preferred the A_4T_4 sequence. DB1947, a smaller compound, showed little to no specificity. DB2240, containing a six-membered central ring with a carbonyl, bound more strongly to all sequences than did DB2239, which contains a similar five-membered central ring. DB2260 bound more strongly or the same as DB2240, the only difference being the replacement of phenyls by pyridines. DB2265, containing a central urea moiety, bound well only to A_4T_4.

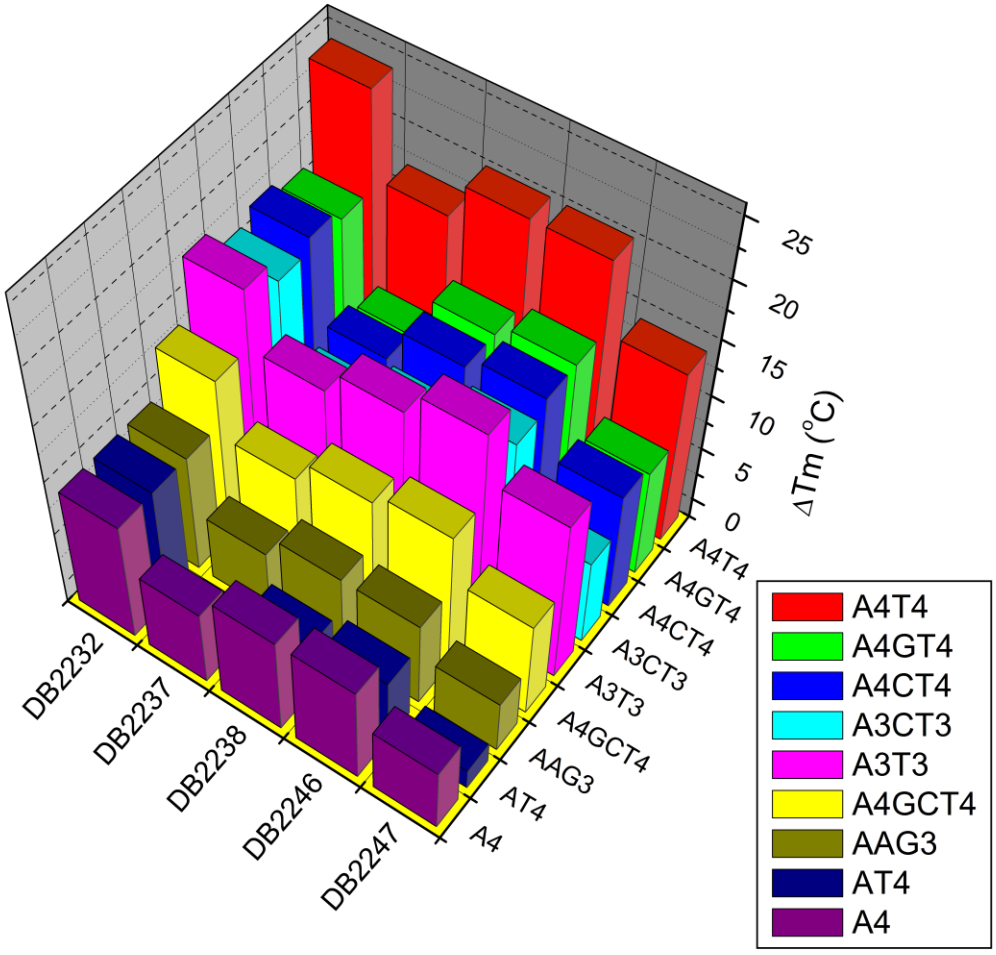


Figure 11. ΔTm results for DB2232 and analogs. Compounds were tested at a 2:1 ratio with 3 µM DNA in CCL or MES buffer (100 mM NaCl) at pH 6.2. DB2232 and analogs predominantly preferred the A_4T_4 and A_3T_3 sequences. Significant binding also occurred with the A_4GT_4, A_4CT_4, A_3CT_3, A_3T_3, and A_4GT_4, sequences. Analogs differed by the location of methyl groups placed on nitrogen atoms of the flanking benzimidazoles. No change in selectivity was observed between the parent compound and analogs. DB2232 remained the strongest overall binder.

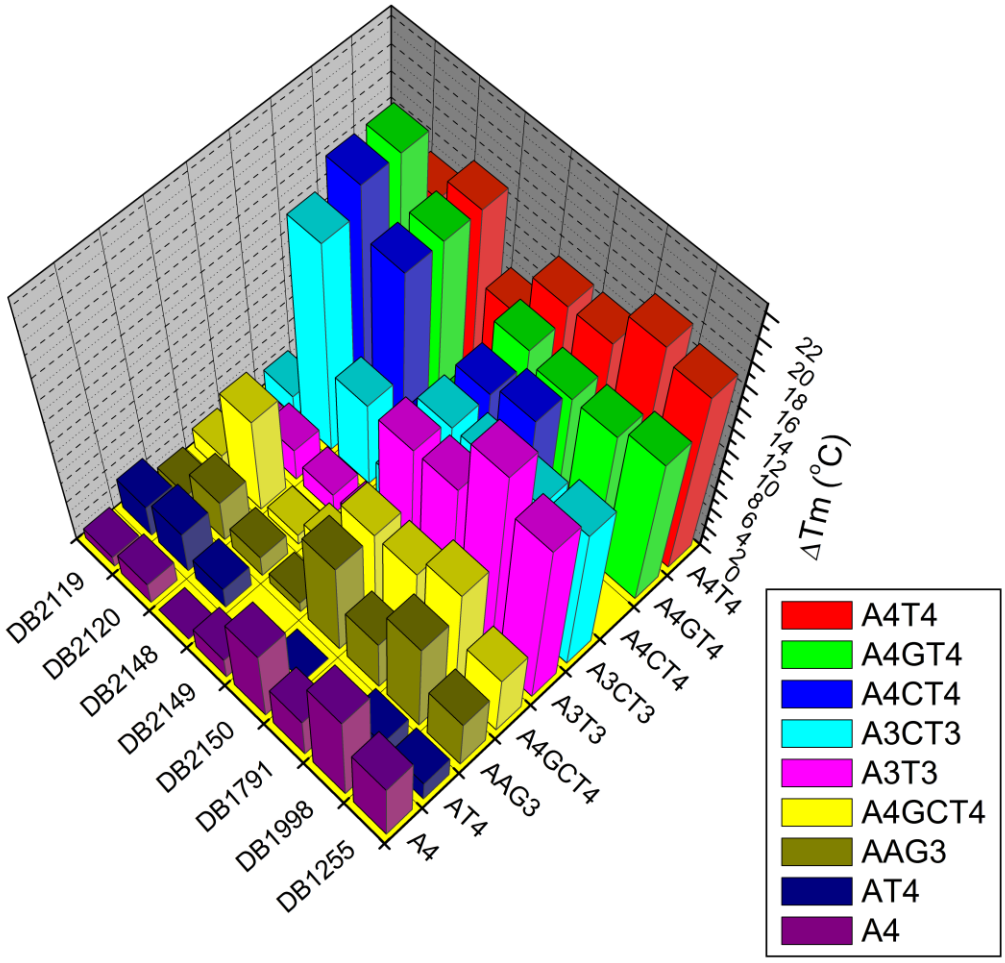


Figure 12. ΔTm results for selected Group II, III, and IV compounds. Compounds were tested at a 2:1 ratio with 3 µM or 1 µM DNA in CCL or MES buffer at pH 6.2. Similar trends in selectivity were observed for DB2148, DB2150, DB1791, DB1998, and DB1255. The A_4T_4 sequence was generally preferred, followed by A_4GT_4, A_4CT_4, A_3CT_3, and A_3T_3. One notable exception is DB2120; this compound prefers A_4GT_4 and A_4CT_4 binding sites to A_4T_4. DB2119 and DB2149 bound to A_4T_4, but were poor binders of the other sequences.

3.3 CIRCULAR DICHROISM

Circular dichrosim experiments were performed to observe induced chirality by ligands binding to DNA. Changes in compound chirality can be seen above about 300 nm, in the compound region of a spectrum, and indicate minor groove binding. Shape changes in DNA can be seen below about 300 nm, in the DNA region of a spectrum, and indicate that binding alters the conformation of DNA in some way. Examples include groove distortion, curvature change, and intercalation.

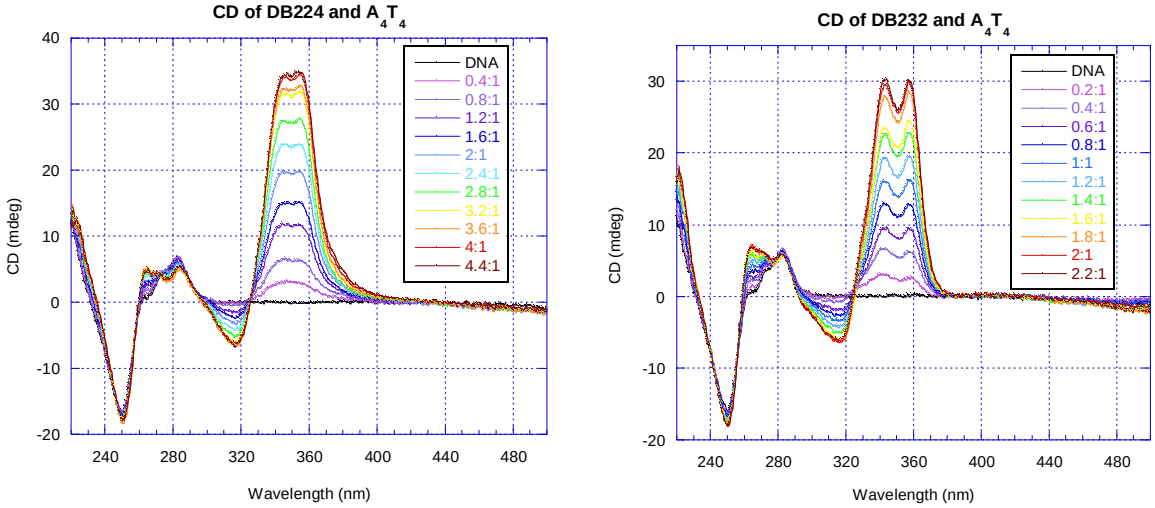


Figure 13. Circular dichroism results for DB224 and DB232 with A_4T_4. Spectra were measured in pH 6.2 CCL buffer (10 mM CCL, 100 mM NaCl, 1 mM EDTA) with a DNA concentration of 3 μM. Similar changes in the compound and DNA regions were observed for DB224 and DB232 with A_4T_4; however, the CD of DB232 was saturated at 2:1 while that of DB224 was saturated at 4:1. DB232 also exhibited a larger induced CD in the DNA region than did DB224, but both spectra indicate some change in DNA structure upon binding. Both compounds show a dramatic induced CD in the compound region, indicating minor groove binding, and saturated at similar intensities despite the different saturating ratios.

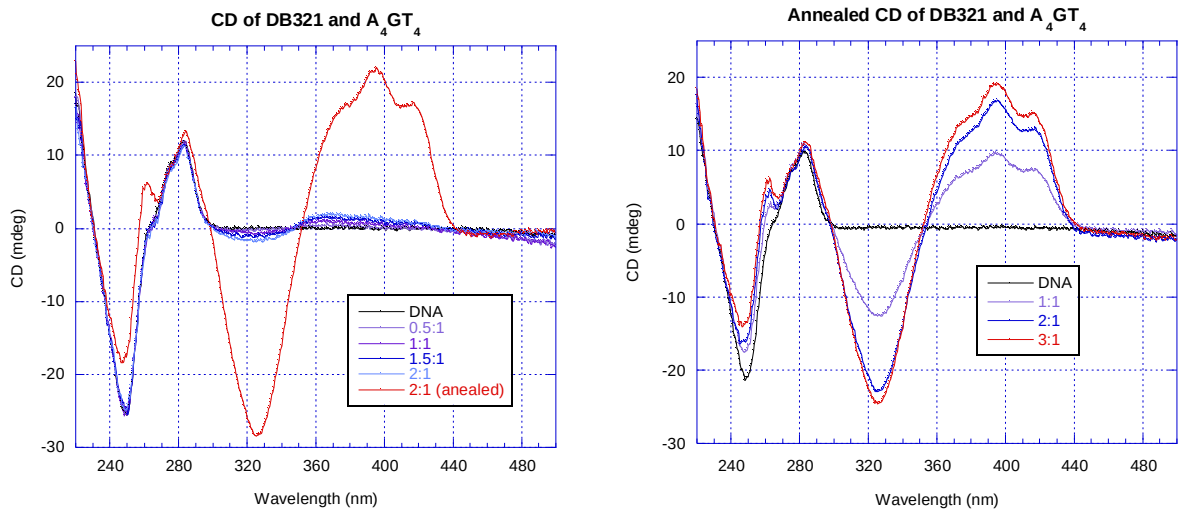


Figure 14. Circular dichroism results for DB321 with A_4GT_4. Spectra were measured in pH 6.2 CCL buffer (10 mM CCL, 100 mM NaCl, 1 mM EDTA) with a DNA concentration of 3 µM.The spectra marked “annealed” are from samples that were removed from the instrument and annealed prior to scanning. Initially, DB321 did not appear to have induced CD for the A_4GT_4 sequence; however, when the samples were annealed prior to being tested, DB321 induced a large CD in the compound and DNA regions of the spectra, indicating minor groove binding and some change to DNA structure upon binding. DB321 saturated the CD spectrum at a ratio of approximately 3:1.

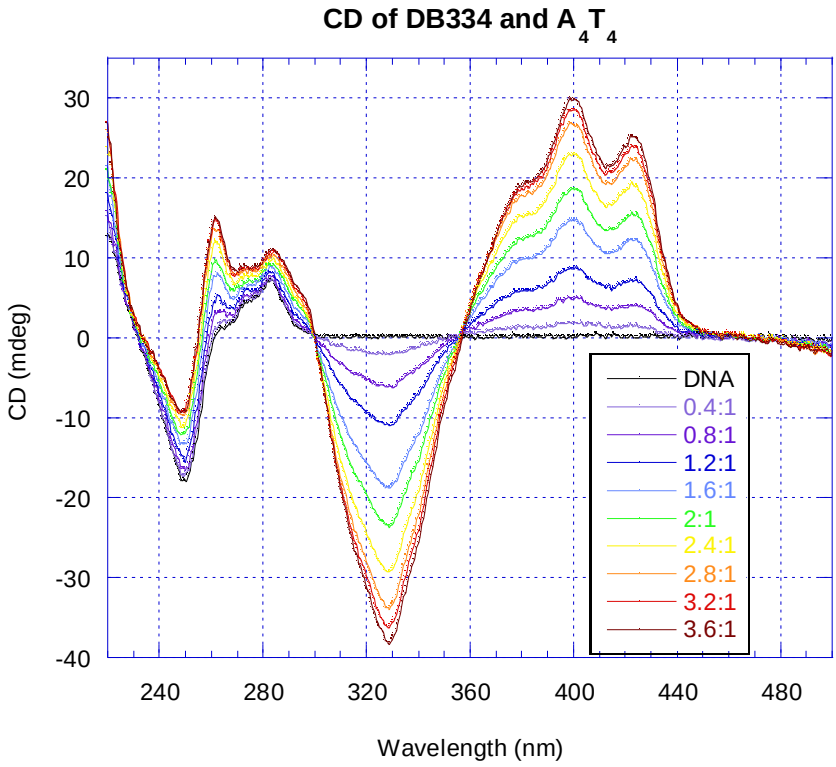


Figure 15. Circular dichroism results for DB334 with A_4T_4. Spectra were measured in pH 6.2 CCL buffer (10 mM CCL, 100 mM NaCl, 1 mM EDTA) with a DNA concentration of 3 μM. The CD of DB334 with A_4T_4 shows a dramatic positive and negative induced CD in the compound region, indicating minor groove binding. A large induced CD is also visible in the DNA region, indicating significant DNA structural changes upon binding. DB334 saturated the CD spectra at a ratio of approximately 3.2:1.

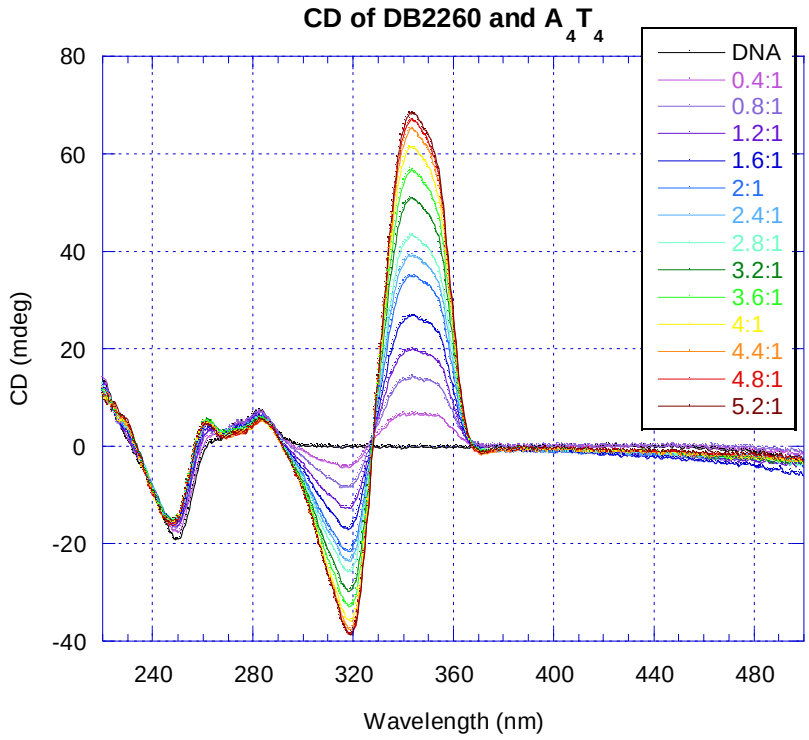


Figure 16. Circular dichroism results for DB2260 with A_4T_4. Spectra were measured in pH 6.2 CCL buffer (10 mM CCL, 100 mM NaCl, 1 mM EDTA) with a DNA concentration of 3 µM. The CD of DB2260 with A_4T_4 shows a large positive and negative induced CD in the compound region, indicating minor groove binding. A smaller induced CD is visible in the DNA region, indicating structural changes to the DNA upon binding. DB2260 saturated the CD spectra at a ratio of approximately 4.4:1.

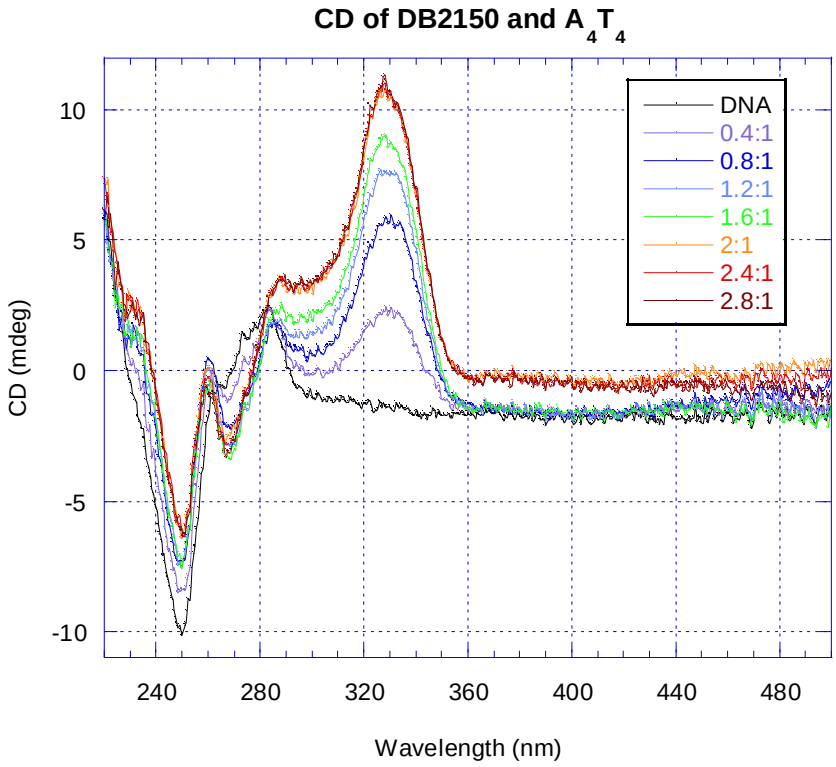


Figure 17. Circular dichroism results for DB2150 with A_4T_4. Spectra were measured in pH 6.2 CCL buffer (10 mM CCL, 100 mM NaCl, 1 mM EDTA) with a DNA concentration of 3 μM. The CD of DB2150 with A_4T_4 shows an induced CD in the compound and DNA regions, indicating minor groove binding and some structural changes to the DNA upon binding. DB2150 saturated the CD spectra at a ratio of approximately 2:1.

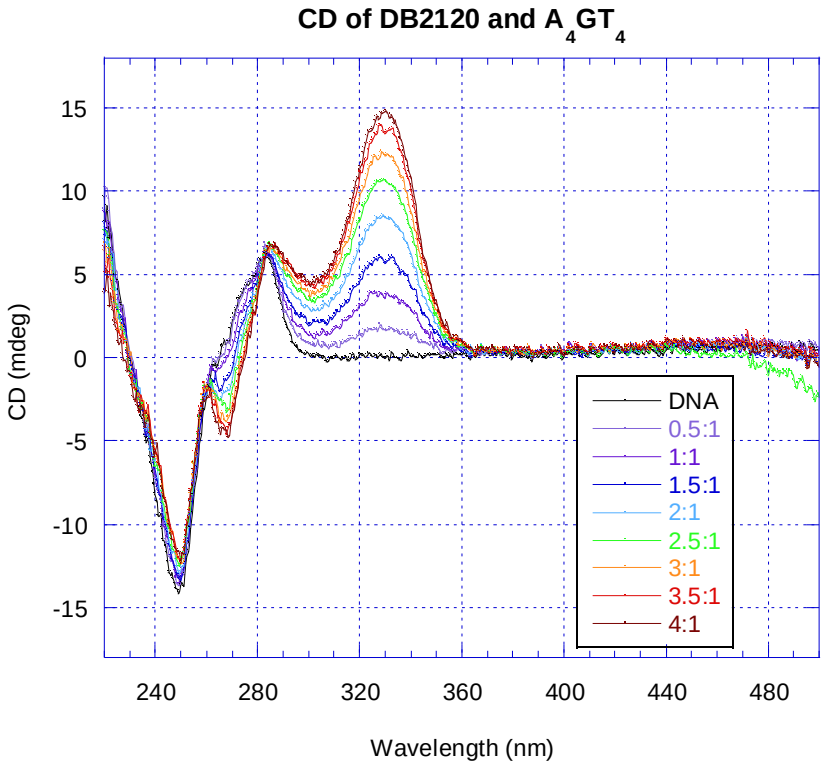


Figure 18. Circular dichroism results for DB2120 with A_4GT_4. Spectra were measured in pH 6.2 CCL buffer (10 mM CCL, 100 mM NaCl, 1 mM EDTA) with a DNA concentration of 3 μM. The CD of DB2120 with A_4GT_4 shows an induced CD in both the compound and DNA region, indicating minor groove binding and structural changes to the DNA upon binding. DB2120 saturated the CD spectra at a ratio of approximately 3.5:1.

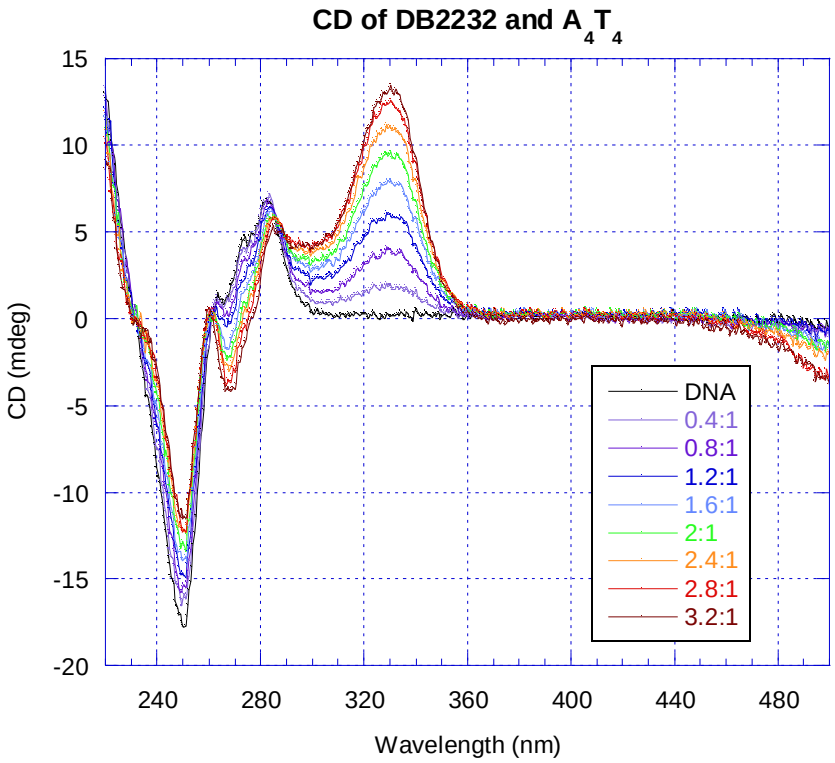


Figure 19. Circular dichroism results for DB2232 with A_4T_4. Spectra were measured in pH 6.2 CCL buffer (10 mM CCL, 100 mM NaCl, 1 mM EDTA) with a DNA concentration of 3 µM. The CD of DB2232 with A_4T_4 shows an induced CD in the compound and DNA regions, indicating minor groove binding and structural changes to the DNA upon binding. DB2232 saturated the CD spectra at a ratio of approximately 2.8:1.

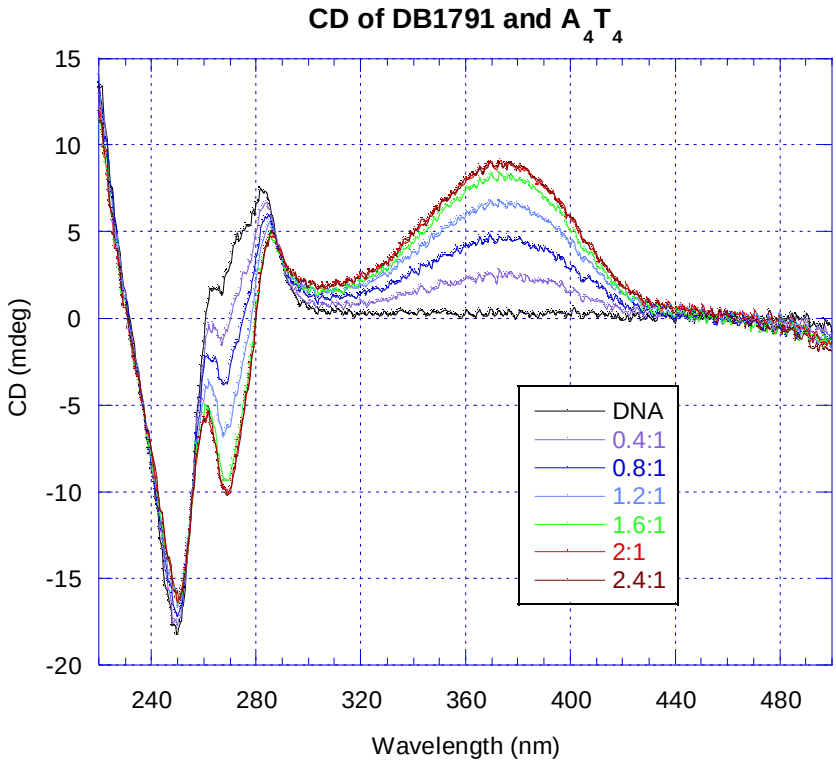


Figure 20. Circular dichroism results for DB1791 with A_4T_4. Spectra were measured in pH 6.2 CCL buffer (10 mM CCL, 100 mM NaCl, 1 mM EDTA) with a DNA concentration of 3 µM. The CD of DB1791 with A_4T_4 shows significant induced CD in the compound region, indicating minor groove binding. A dramatic induced CD is visible in the DNA region of the spectrum, indicating significant DNA structural changes upon binding. DB1791 saturated the spectra at a ratio of approximately 1.6:1.

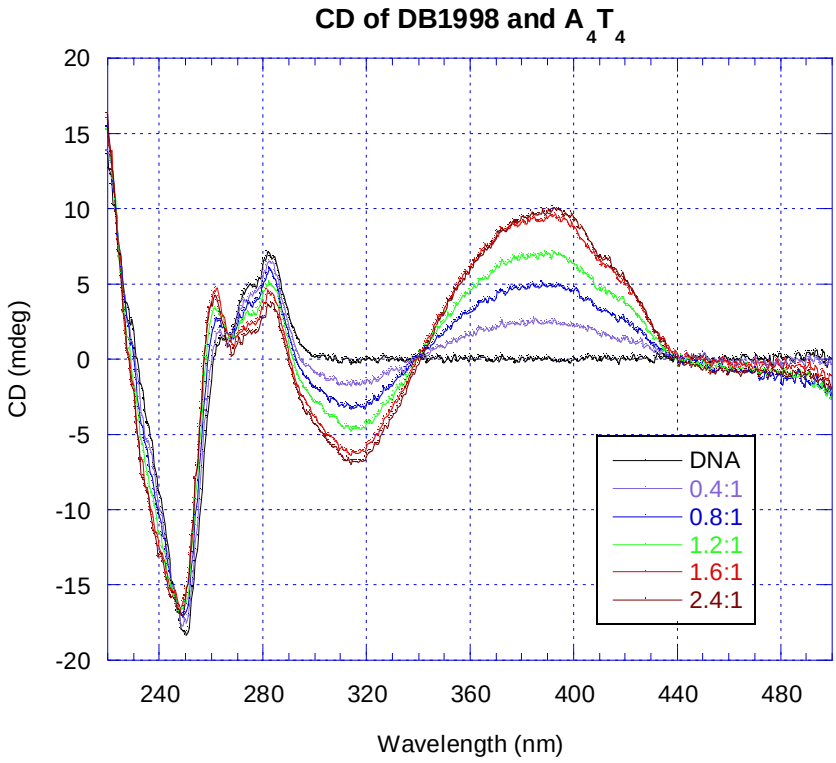


Figure 21. Circular dichroism results for DB1998 with A_4T_4. Spectra were measured in pH 6.2 CCL buffer (10 mM CCL, 100 mM NaCl, 1 mM EDTA) with a DNA concentration of 3 µM. The CD of DB1998 with A_4T_4 shows both positive and negative induced CD in the compound region, indicating minor groove binding. A smaller induced CD is visible in the DNA region, indicating some structural change to the DNA upon binding. DB1998 saturated the spectra at a ratio of approximately 1.6:1.

Table 3. Approximate saturating ratios for CD experiments. Spectra were measured in pH 6.2 CCL buffer (10 mM CCL, 100 mM NaCl, 1 mM EDTA) with a DNA concentration of 3 μM. DB232 saturated at a much lower ratio than did its analog DB224, which suggests that DB232 is a stronger binding compound. DB321 and its analog, DB334, saturated the spectra of different sequences at similar ratios; however, the samples of DB321 were annealed between each titration. DB2150, DB1791, and DB1998 all saturated at or below approximately a 2:1 ratio, suggesting that these compound are relatively strong binders. DB2120 saturated at about 3.5:1, DB2232 saturated at about 2.8:1, and DB2260 saturated at a high ratio of approximately 4.4:1.

Compound	Sequence	Appx. Saturating Ratio
DB224	A_4T_4	3.8-4.2
DB232	A_4T_4	1.8-2.2
DB321	A_4GT_4	2.8-3.2
DB334	A_4T_4	3-3.4
DB2260	A_4T_4	4.2-4.6
DB2150	A_4T_4	1.8-2.2
DB2120	A_4GT_4	3.3-3.7
DB2232	A_4T_4	2.6-3
DB1791	A_4T_4	1.4-1.8
DB1998	A_4T_4	1.4-1.8

3.4 MASS SPECTROMETRY

Mass spectrometry experiments were performed to determine the stoichiometry of binding. Multiple DNA sequences were tested together with a single compound for speed and qualitative assessment of relative binding.

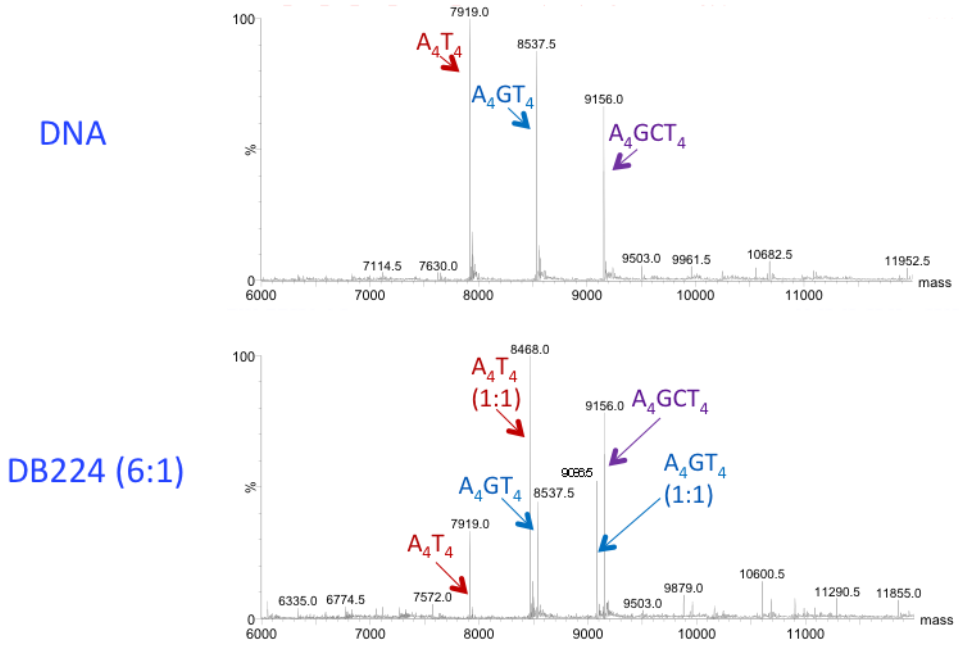


Figure 22. Mass spectrometry results for DB224. Samples containing 5 μM of each oligonucleotide were tested in 150 mM ammonium acetate buffer with a capillary voltage of 2.5 kV, cone voltage of 30 V, source block temperature of 70° C, desolvation temperature of 100° C, nitrogen as the nebulizing and drying gas, and direct injection at 5 μL/min. Using MassLynx 4.1 software, the observed mass of each equivalent species was multiplied by its charge and the areas were added to give a single peak at the "parent" mass. Ratios are reported as [compound]:[each oligonucleotide]; samples were prepared from the same pre-mixed DNA stock solution. DB224 bound as a monomer to A_4T_4 and A_4GT_4. No complex is visible for A_4GCT_4, which is in agreement with the ΔTm data. This result shows that mass spectrometry is an excellent method to determine binding stoichiometry, and can also be used as a qualitative comparison for major differences in binding affinity.

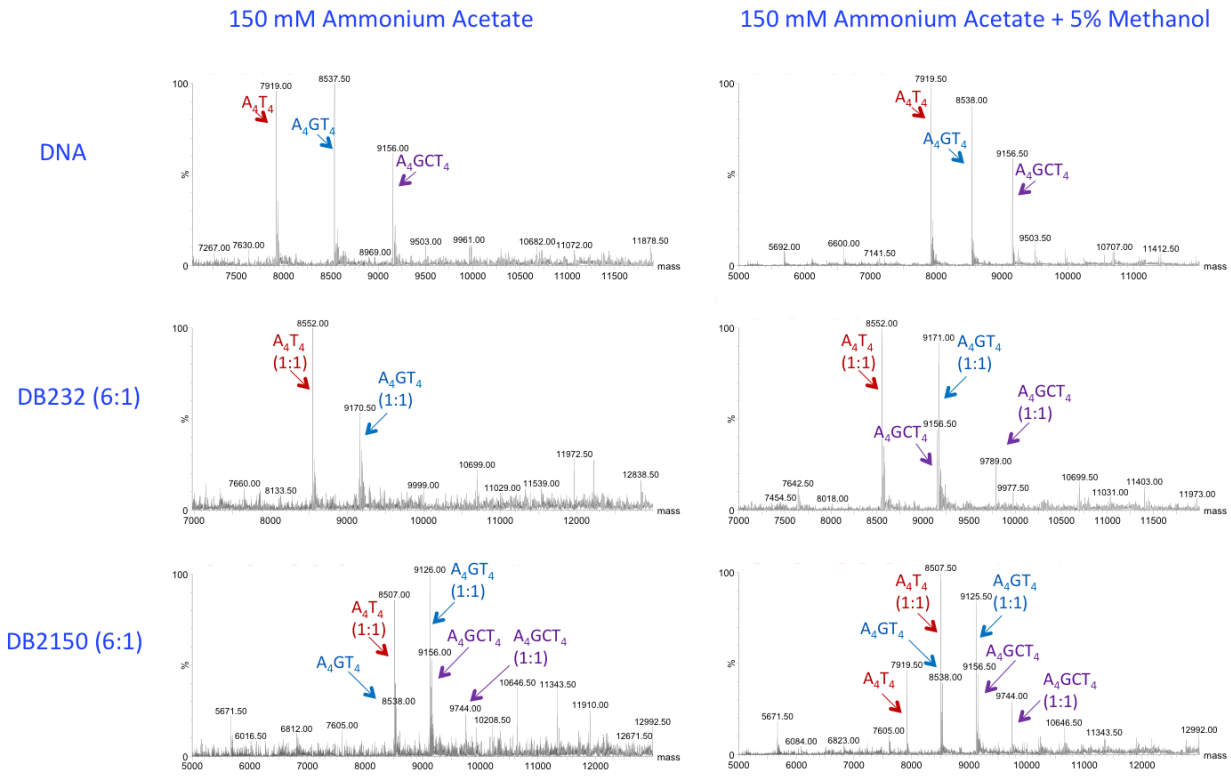
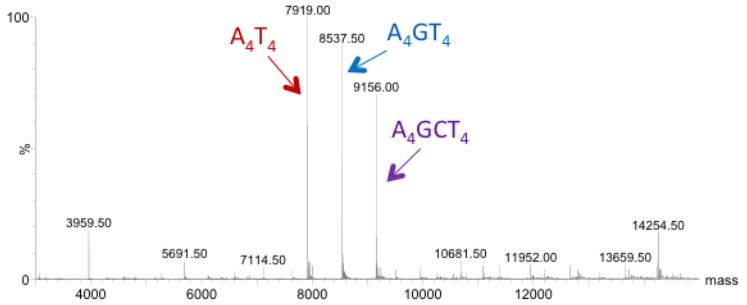
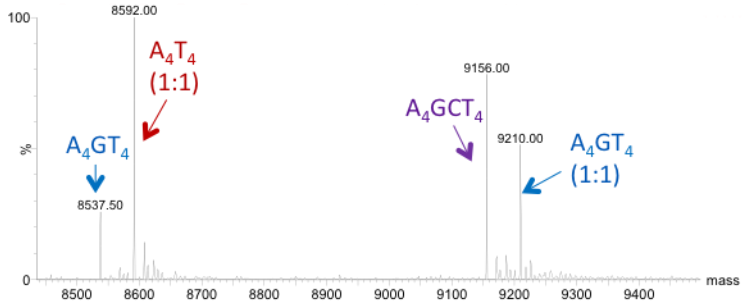


Figure 23. Mass spectrometry with and without methanol. Samples containing 5 μM of each oligonucleotide were tested in 150 mM ammonium acetate with and without 5% methanol (v/v) buffer, with a capillary voltage of 2.5 kV, cone voltage of 30 V, source block temperature of 70° C, desolvation temperature of 100° C, nitrogen as the nebulizing and drying gas, and direct injection at 5 μL/min. Using MassLynx 4.1 software, the observed mass of each equivalent species was multiplied by its charge and the areas were added to give a single peak at the "parent" mass. Ratios are reported as [compound]:[each oligonucleotide]; all samples were prepared from the same pre-mixed DNA stock solution. DB232 bound as a monomer to A_4T_4, A_4GT_4, and A_4GCT_4. DB2150 also bound as a monomer to A_4T_4, A_4GT_4, and A_4GCT_4. In all cases, the samples tested with 5% methanol resulted in a better signal-to-noise ratio. The complex of DB232 with A_4GCT_4 was only visible in the spectrum taken with methanol.

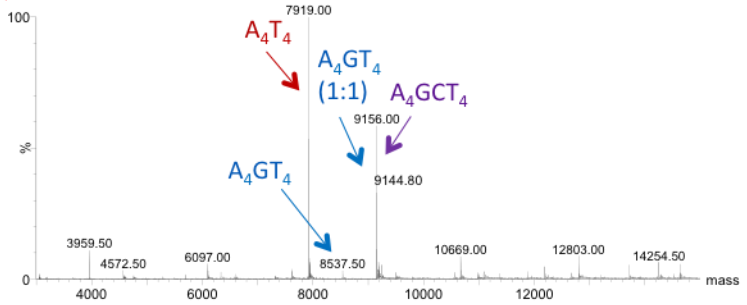
DNA



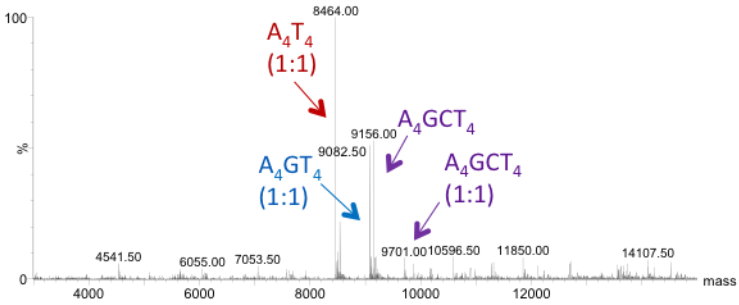
DB334 (6:1)



DB2120 (6:1)



DB2232 (6:1)



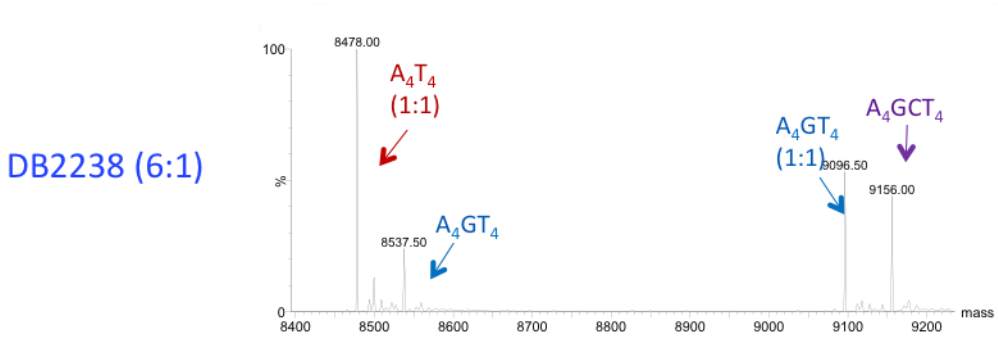


Figure 24. Mass spectrometry results for DB334, DB2120, DB2232, and DB2238. Samples containing 5 μM of each oligonucleotide were tested in 150 mM ammonium acetate/5% methanol (v/v) buffer with a capillary voltage of 2.5 kV, cone voltage of 30 V, source block temperature of 70° C, desolvation temperature of 100° C, nitrogen as the nebulizing and drying gas, and direct injection at 5 μL/min. Using MassLynx 4.1 software, the observed mass of each equivalent species was multiplied by its charge and the areas were added to give a single peak at the "parent" mass. Ratios are reported as [compound]:[each oligonucleotide]; all samples were prepared from the same pre-mixed DNA stock solution. DB334 bound as a monomer to A_4T_4 and A_4GT_4, but no complex was visible with A_4GCT_4. DB2120 bound as a monomer to A_4GT_4, but no complex was visible with A_4T_4 or A_4GCT_4. This is in agreement with the selectivity of DB2120 shown via thermal melting experiments. DB2232 bound as a monomer to A_4T_4, A_4GT_4, and A_4GCT_4. DB2238 bound as a monomer to A_4GT_4 and A_4T_4, but no complex was visible with A_4GCT_4

Table 4: Summary of mass spectrometry results. Samples containing 5 μM of each oligonucleotide and a 6:1 ratio of [compound]:[each oligonucleotide] were tested in 150 mM ammonium acetate with or without 5% methanol (v/v) buffer. “1:1” indicates monomeric binding; “0 “ indicates that no complex was formed. Monomeric complexes were observed for DB224 with A_4T_4 and A_4GT_4; for DB232 with A_4T_4, A_4GT_4, and A_4GCT_4; for DB334 with A_4T_4 and A_4GT_4; for DB2120 with A_4GT_4; for DB2150 with A_4T_4, A_4GT_4, and A_4GCT_4; for DB2232 with A_4T_4, A_4GT_4, and A_4GCT_4; and for DB2238 with A_4T_4 and A_4GT_4. No 2:1 complexes were observed.

	A_4T_4	**A_4GT_4**	**A_4GCT_4**
DB224	1:1	1:1	0
DB232	1:1	1:1	1:1
DB334	1:1	1:1	0
DB2120	0	1:1	0
DB2150	1:1	1:1	1:1
DB2232	1:1	1:1	1:1
DB2238	1:1	1:1	0

3.5 SURFACE PLASMON RESONANCE

Surface plasmon resonance studies were performed for determination of binding constants. Several compounds were tested via SPR with HEPES buffer; however, these large compounds resulted in significant streptavidin-surface binding. MES buffer gave better results for a few cases, and was used instead for all surface plasmon resonance studies. Many of the compounds in Groups I and II were tested in this manner, but most resulted in unpresentable sensorgams due to streptavidin-surface binding, aggregation, and general “sticky” compound properties.

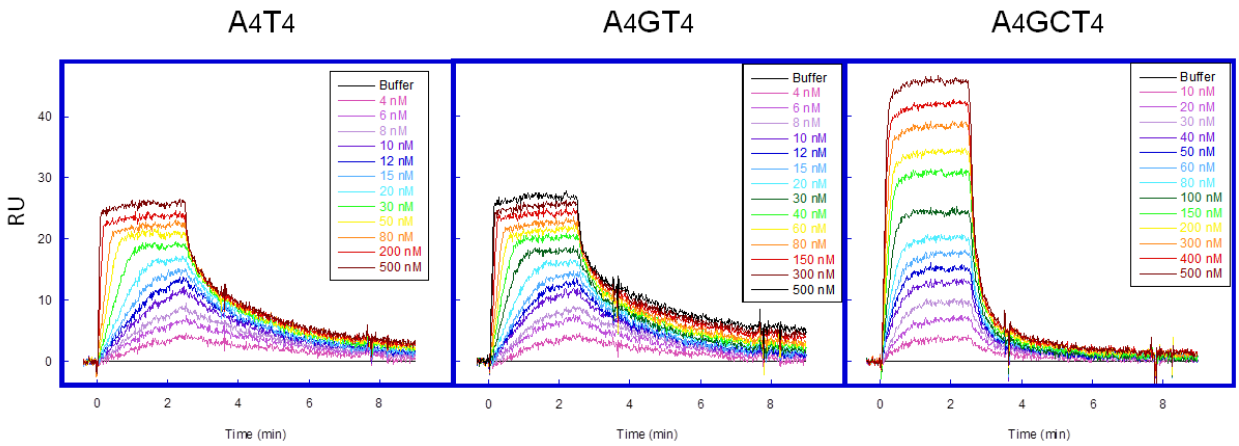


Figure 25. SPR Sensorgrams for DB2150. Samples were prepared and tested in pH 6.2 MES buffer (0.005% P20, v/v) at various ratios from 5 nM to 500 nM. Data was collected using a Biacore 2000 optical biosensor instrument.

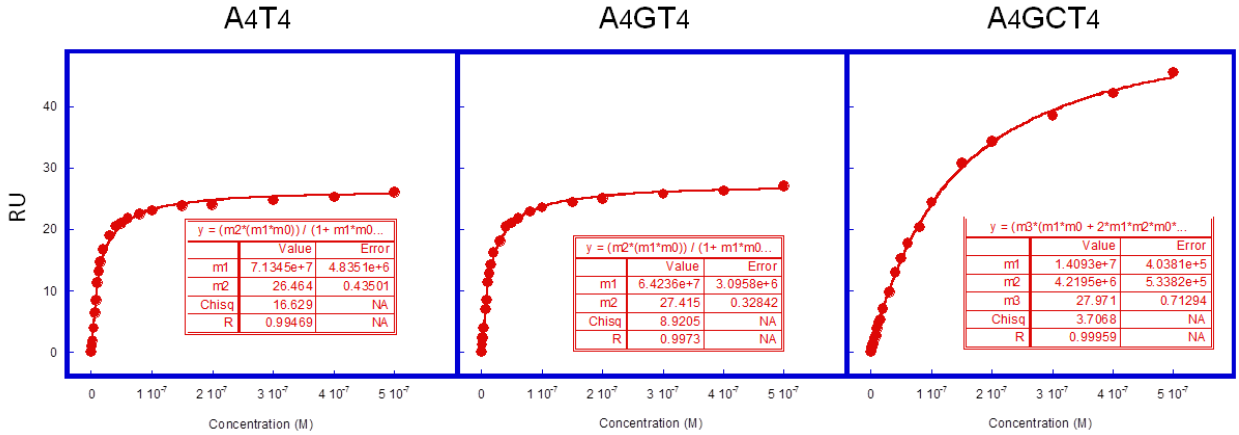


Figure 26. SPR Steady-State Fits for DB2150. Samples were prepared and tested in pH 6.2 MES buffer (0.005% P20, v/v) at various ratios from 5 nM to 500 nM. Data was collected using a Biacore 2000 optical biosensor instrument. DB2150 bound with similar affinity to A_4T_4, A_4GT_4, and A_4GCT_4, with binding constants of 7.1 x 10^7, 6.4 x 10^7, and average binding constant of 9.1 x 10^6 respectively. The dimeric binding to A_4GCT_4 showed negative cooperatively. The stoichiometric values used to for data fitting were based on the predicted maximum RU values for the different sequences.

MOLECULAR MODELING

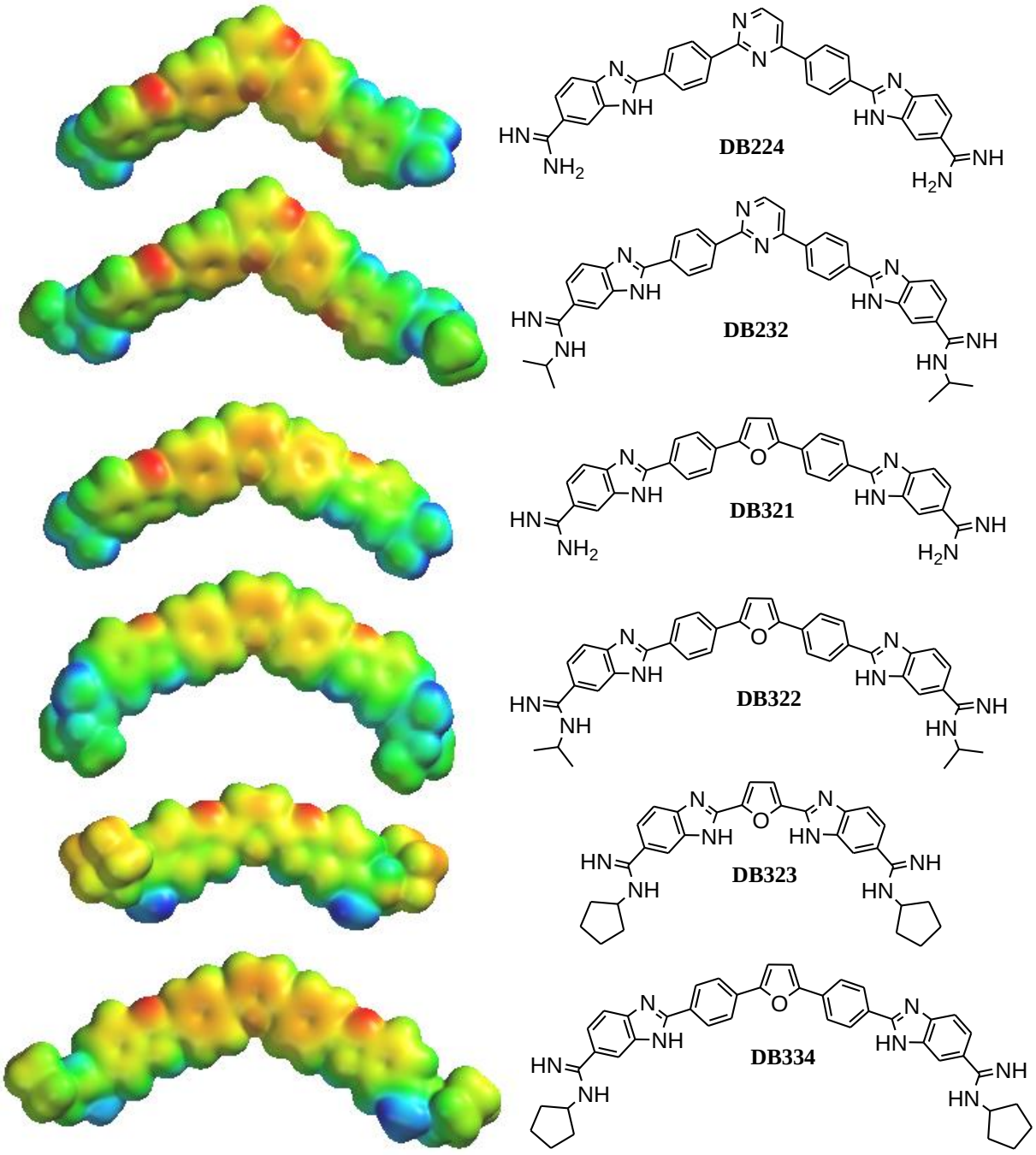


Figure 27. Electrostatic potential maps for minimum energy conformers of DB224, DB321, and analogs. Maps were generated by Spartan '04 using a Hartree-Fock model with a 6-31G* basis set to determine the conformation with E_{min}. Red indicates regions of negative electrostatic potential, green indicates hydrophobic regions, and blue indicates regions of positive electrostatic potential. Note the increased hydrophobic surfaces in DB232, DB322, and DB334.

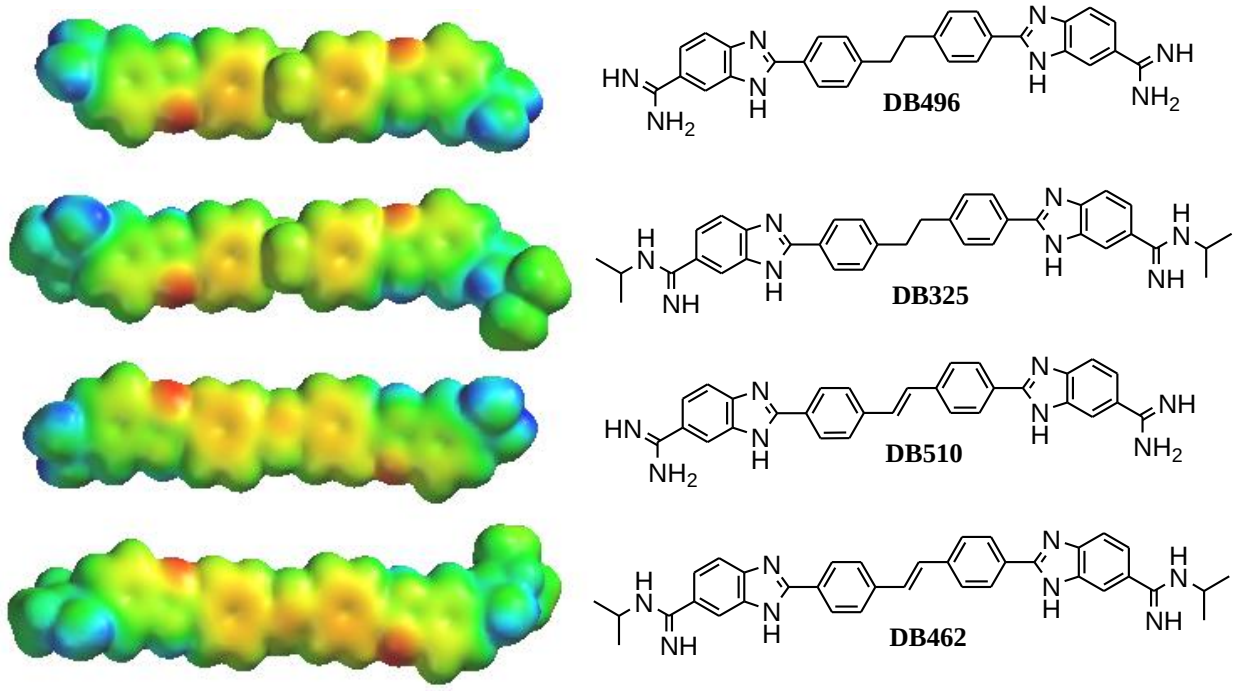


Figure 28. Electrostatic potential maps for minimum energy conformers of DB496, DB510, and analogs. Maps were generated by Spartan ’04 using a Hartree-Fock model with a 6-31G* basis set to determine the conformation with E_{min}. Red indicates regions of negative electrostatic potential, green indicates hydrophobic regions, and blue indicates regions of positive electrostatic potential. Note the increased hydrophobic surfaces in DB325 and DB462.

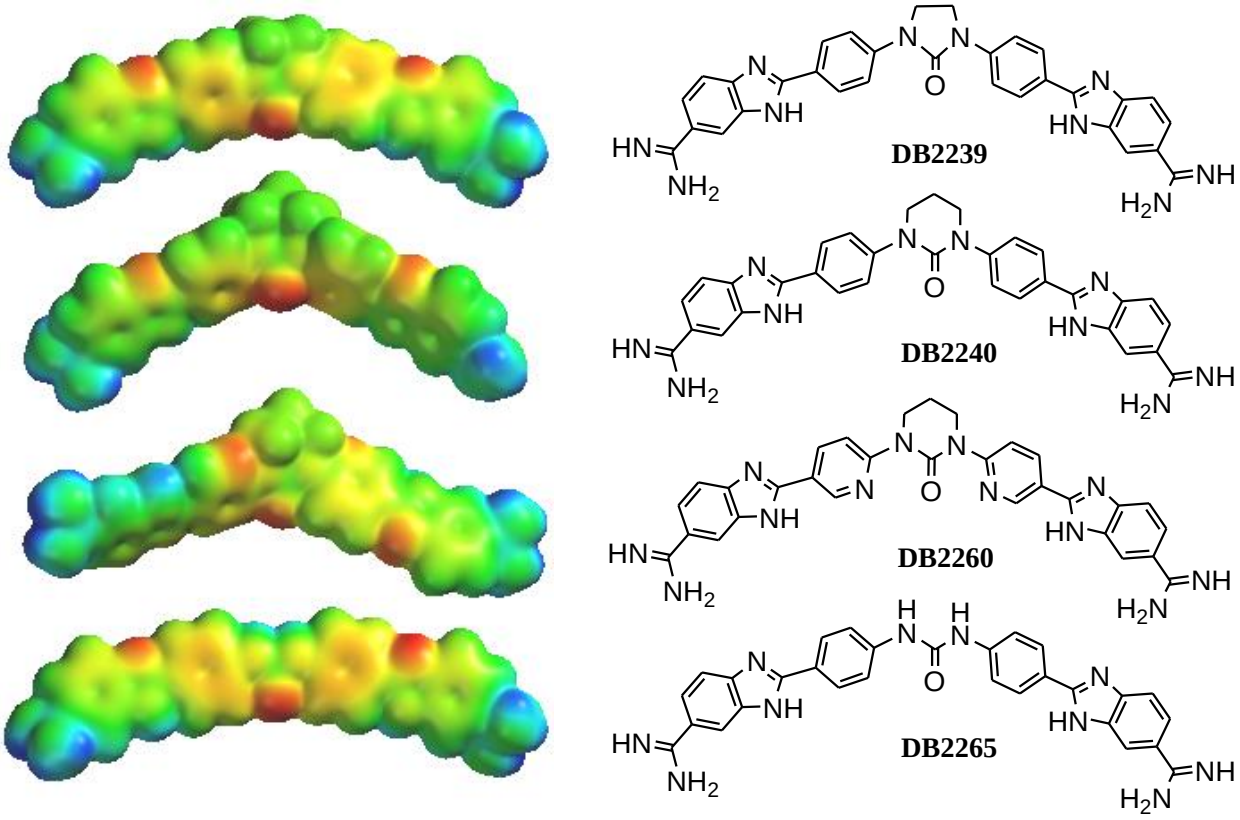


Figure 29. Electrostatic potential maps for minimum energy conformers of selected Group I compounds. Maps were generated by Spartan '04 using a Hartree-Fock model with a 6-31G* basis set to determine the conformation with E_{min}. Red indicates regions of negative electrostatic potential, green indicates hydrophobic regions, and blue indicates regions of positive electrostatic potential. The geometry of DB2240 appears drastically changed when compared to DB2239. DB2260 is very similar to DB2240, but DB2265 is almost linear.

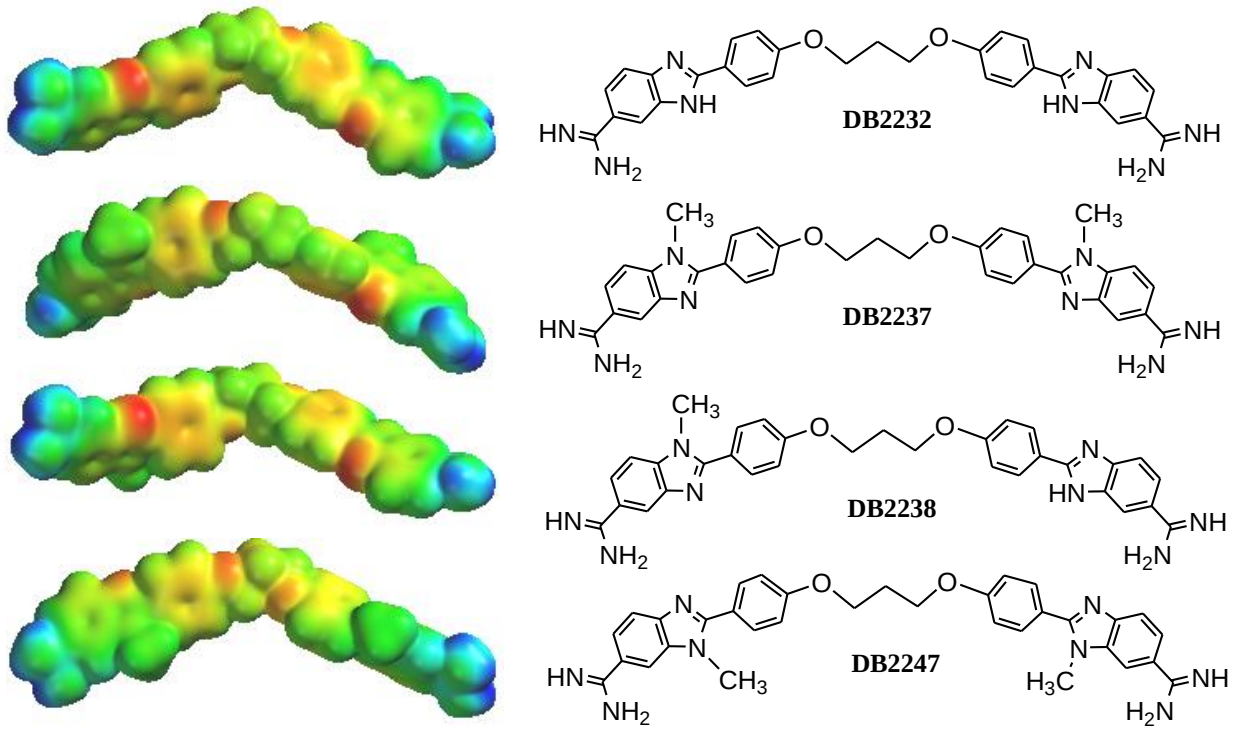


Figure 31. Electrostatic potential maps for minimum energy conformers of DB2232 and analogs. Maps were generated by Spartan ’04 using a Hartree-Fock model with a 6-31G* basis set to determine the conformation with E_{min}. Red indicates regions of negative electrostatic potential, green indicates hydrophobic regions, and blue indicates regions of positive electrostatic potential. Note the steric hindrance to binding caused by the introduction of methyl groups onto the flanking benzimidazoles.

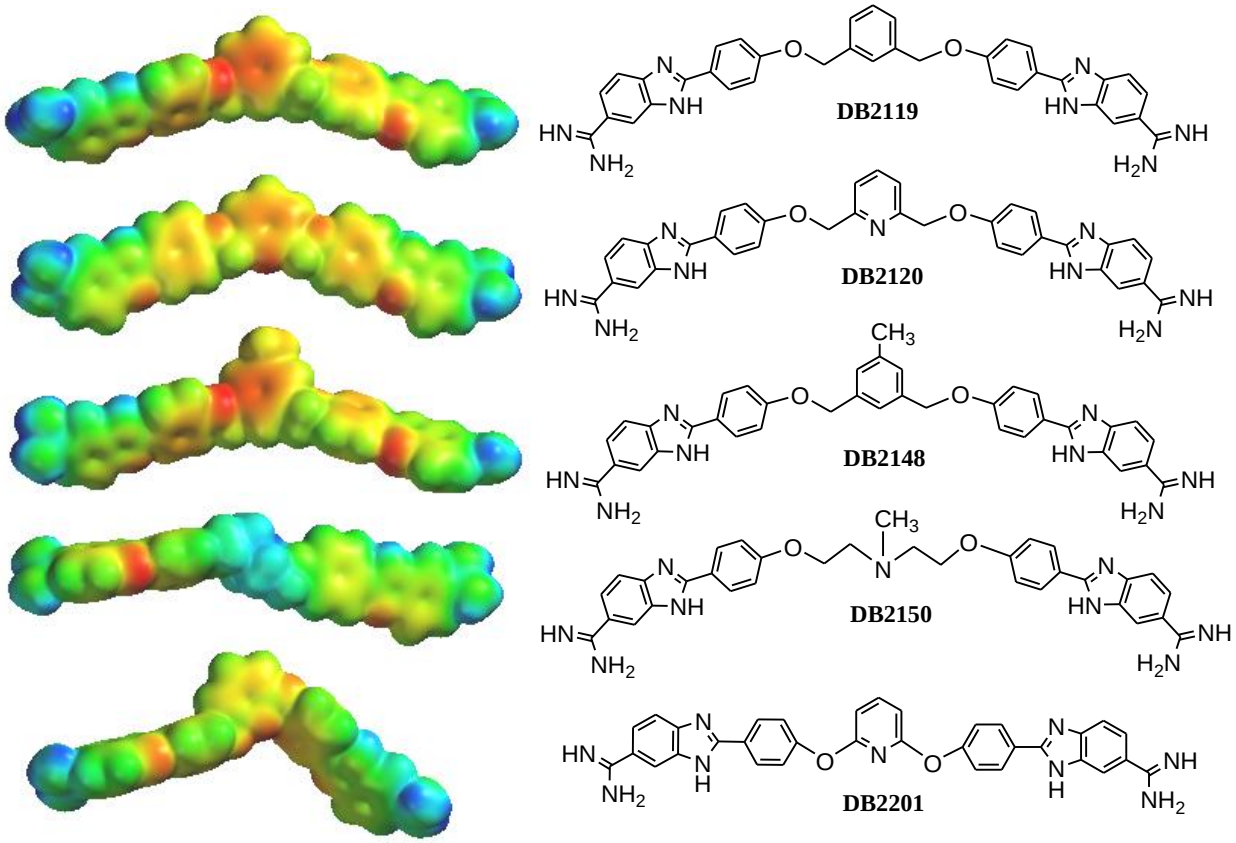


Figure 32. Electrostatic potential maps for minimum energy conformers of selected Group II compounds Maps were generated by Spartan '04 using a Hartree-Fock model with a 6-31G* basis set to determine the conformation with E_{min}. Red indicates regions of negative electrostatic potential, green indicates hydrophobic regions, and blue indicates regions of positive electrostatic potential. DB2119 and DB2120 have different minimized conformations despite their structural similarity. DB2148 and DB2119 both show the central ring twisted out of the plane created by the flanking groups. DB2150 was modeled as a trication, and as such has enhanced electropositive potential. The model of DB2201 shows a conformation inconsistent with typical minor groove binders due to the geometry of the ether linkages. This conformation is in stark contrast to that of DB2120, which differs only by the inclusion of two methylene groups.

4 DISCUSSION

4.1 DNA Microstructure

Poly-dA oligonucleotides such as A_4T_4 are more highly curved than mixed sequences.[4-5] The minor groove narrows relative to sequences containing GC BPs,[16] but less so for sequences containing multiple AT steps.[6] The bases also take on a propeller twist in A-tracts.[4, 16] The narrowed minor groove allows stronger contacts with groove-binders and greater negative electrostatic potential.[4-6] The combination of these forces accounts for the overarching binding preferences evident most clearly in Group I compounds.

Since mixed A-T sequences take on A-tract characteristics to a lesser extent, compounds with a high affinity for A_4T_4 bind to $(AT)_4$ with a lower K and a less favorable ΔG. Alternating A-T sequences also tend to have overwound helices compared to mixed sequence B-DNA, which has a helix repeat of about 10.5 BPs per turn.[4-5, 17] In the case of poly-dAT or poly-dAAT sequences, a D-DNA has been observed, which has a helix repeat of about 8 BPs per turn; but alternating A-T sequences are quite polymorphic. None of the compounds tested were able to bind to the $(AT)_4$ sequence with much success. Introduction of GC base pairs in the middle of the binding region causes the DNA to adopt a wider minor groove and a more regular curvature, a bad fit for conformationally locked compounds. A longer linker could solve this problem; however, some structural rigidity should be maintained if the compound is to be selective.

The ΔTm results indicate that the sequence A_4GCT_4 is quite different from A_4GT_4 and A_4CT_4, which have many similarities with A_4T_4. Two GC BPs effectively disrupt the A-tract curvature and present a large steric hindrance to binding in the form of exocyclic amines in the minor groove. Most of the compounds tested did not bind well to the A_4 single-site control sequence in the minor groove, which indicates that the longer, two-site binding sequence is required. The $(AAG)_3$ site was also challenging to recognize. Having only two AT BPs on either side of a GC BP decreased minor groove electrostatic potential and binding affinity of compounds. Once a successful paradigm is developed for GC recognition, such a complex sequence should be recognizable.

4.2 Group I Compounds

Minor groove binders typically have a characteristic crescent shape,[4] a motif that holds true for many of the examined compounds. Evident from the molecular models (Figures 28-32), many of the relatively rigid compounds in Group I are slightly more curved than most. Their limited accessible conformations cause tighter than normal curvatures, providing the physical basis for these compounds' specificity of binding. Binding of any ligand to the minor groove is expected to decrease the number of available conformations accessible to the DNA. It follows that a more conformationally limited ligand will result in a more conformationally limited complex, and that a more rigid compound could mean a greater entropic loss by DNA upon binding. However, this conformational entropic loss should be somewhat offset by the release of water from the minor groove.

Compounds in Group I which were able to significantly bind to A_4GT_4 and A_4CT_4 (ΔTm > 10 ºC) did so with a similar or lessened affinity compared to A_4T_4, indicating that the hydrogen bond acceptors in the linker region are most likely not binding to the amine of guanine. Interestingly, the larger compounds from Group I, such as DB232 and DB321, showed very low affinity for smaller sequences such as A_4. Again, this selectivity suggests the two-site binding mode for these compounds. Also, since DB321 could bind well to A_3CT_3 but not to A_3T_3, this particular compound may require a binding site at least seven base pairs long. DB2147 bound well only to the A_4T_4 sequence. This compound is extremely curved, so the inclusion of even one GC base pair may have relaxed the minor groove curvature enough to make DB2147 a bad fit.

The terminal amidine-isopropyl or amidine-cyclopentane analogs (DB232, DB322, DB235, DB334, and DB462) bound with a greater thermal melting point elevation than their parent, terminal-amidine compounds (DB224, DB321, DB496, and DB510). These analogs, containing non-polar amidine extensions, experience increased binding in a non-specific manner. Such an inclusion would likely do the same for many minor groove binders. Shown in Figure 13, the CD of DB232 with A_4T_4 saturated at a much lower ratio than did DB224 (Table 3), again, suggesting increased binding affinity. Although DB510 is extremely close in structure to DB462, it showed very little change in thermal melting temperature for any sequence, indicating that the flanking isopropyl groups are essential to the binding of this structure. Likewise, DB496 showed almost no binding affinity, whereas DB325 dramatically increased the thermal melting temperature of many sequences. This increased DNA affinity is most likely due to van der Waals interactions between the non-polar amidine extension and the

deoxyribose of the DNA. The increased van der Waals surfaces can clearly be seen from the electrostatic potential maps. The molecular models also suggest that the positive charge on the amidine groups becomes partially dispersed through the non-polar extensions. Such a change may decrease the Gibb's energy necessary for compound dissolution for DB224 and DB321, as the interaction between the positive terminals and the central hydrogen bond acceptor would be decreased. Indeed, the circular dichrosim results for DB321 (Figure 14) and DB334 (Figure 15) as well as the thermal melting results for DB224 and DB232 with A_4T_4 (Figure 8) suggest that the inclusion of terminal cyclopentane groups may have decreased stacking effects. Specifically, the inclusion of terminal cyclopentane groups in DB334 eliminated the need to anneal between CD titrations, shown for DB321 in Figure 14.

The molecular modeling results suggest that many of the linkers, designed to accept a hydrogen bond from guanine, may actually be shielded within the molecule and inaccessible due to the flanking phenyl rings. Because of this structural feature, DB2239 and DB2240 were synthesized and are notable exceptions, with the carbonyl group protruding to a more easily accessible location. DB2240, which contains a six-membered central ring, is a stronger binding compound than DB2239, which contains a five-membered central ring. The models show that the change from a six-membered to a five-membered central ring has a profound effect on the geometry of the ligand, with the central ring of DB2239 lying in the plane created by the flanking phenyl-benzimidazole groups, and the central ring of DB2240 twisting out of the plane created by the flanking phenyl-benzimidazole groups. This structural difference can clearly explain the difference in melting point elevation. Paramount to GC recognition will be

the accessibility of a hydrogen bond acceptor group. A comparison between DB323 and DB334 indicates that the phenyl ring does contribute significantly to binding for all sequences except A_4; however, this ring probably interacts in a weakly specific manner to AT BPs, and presents a structural feature that may be changed for greater specificity. By replacing the phenyl with a smaller group, such as an alkyl ether, ethyl, or amide, other hydrogen bond accepting groups could become available. The replacement of the phenyl groups in DB2240 by pyridine groups in DB2260 causes an increase in binding, notably to the A_4GCT_4 sequence. This may be due to decreased steric hindrance, as adjacent hydrogens of the phenyl groups which face the minor groove are effectively eliminated with this substitution.

4.3 Group II Compounds

Thermal melting results suggest that Group II compounds generally do not bind as well as Group I compounds to the favored sequences. Perhaps due to increased flexibility and ability to fold back on themselves, many Group II compounds also appear to be more able to accommodate the single site sequence when compared to the two-site sequences containing one GC BP. For example, DB2150, DB2232, DB2237, DB2238, and DB2246 all elevate the thermal melting point of the A_4GT_4 sequence only 4-7 ºC more than the A_4 sequence, compared to DB224, DB232, DB321, DB322, and DB334, which elevate the A_4GT_4 sequence 15-21 ºC more than the A_4 sequence. The ability to take on a secondary binding mode where each AT site accommodates one folded-back compound, if indeed present, presents a serious challenge to selectivity.

DB2148 is one interesting exception to this lack of selectivity. Perhaps due to the electrostatic potential of the DNA sequences, this compound does not bind well to any sequence other than A_4T_4, A_4GT_4 and A_4CT_4. DB2119 exhibits similar selectivity; however, the analog DB2120 is strikingly selective for sequences with a single G or C, elevating the thermal melting point of A_4GT_4 and A_4CT_4 over 21 °C. The molecular modeling results for DB2119 and DB2120 (Figure 32) indicate significantly different minimized structures, with the central phenyl of DB2119 twisting out of the plane created by the flanking phenyl-benzimidazole groups, and the central pyridine of DB2120 lying in plane with the flanking phenyl-benzimidazole groups. This is likely a contributing factor to their differences in selectivity. Although DB2120 shows significant binding to the A_4T_4 sequence (ΔTm of 15 °C), this compound is an excellent lead for the development of linkers capable of recognizing one or two GC BPs. DB2120 combines a central aromatic ring with flexible linkers, which perhaps allow the adjacent hydrogens of the phenyl groups enough room to prevent steric clashes.

DB2150 binds with a similar affinity to A_4T_4 and A_4GT_4, and with a slightly lessened binding affinity to A_4GCT_4. These SPR results are reinforced by the ΔTm results; however, the mass spectrometry and SPR results are not in good agreement as to the stoichiometry of the DB2150 and A_4GCT_4 complex (Figures 23 and 26). Mass spectrometry indicates a 1:1 complex, while SPR indicates a 2:1 complex. This discrepancy can be explained by the negative cooperativity shown via SPR. Since the first binding constant is significantly higher than the second, the second molecule may be binding through non-specific interactions. DB2150 is assumed to be a trication in solution, and the dimeric complex may be too unstable to enter the gas phase, even

using ESI mass spectrometry. The monomeric complex would be more stable, thus visible by mass spectrometry.

The N-CH_3 substituted benzimidazole analogs of DB2232 were developed in an effort to understand the effects of the protonation states of the benzimidazoles on affinity and selectivity. The results suggest no change in selectivity, only decreases in binding most likely due to steric effects and a loss in the ability to act as hydrogen bond donors. Interestingly, the mass spectrometry results for DB2232 (Figure 24) show only monomeric complexes with A_4T_4, A_4GT_4, and A_4GCT_4. This is in stark contrast with a previously reported dimer that forms with the sequence AATTGCAATT.[18] The proposed binding model involved RT546 (DB2232) folding back on itself and binding one folded molecule to each AATT site. Because the minor groove of AATT sites is wider relative to A_4 sites, DB2232 may not be able to fit in such a narrowed minor groove when folded back. The discrepancy between the monomeric and dimeric results may arise from specific properties inherent with flanking A_4 sites; perhaps this tightened minor groove must be wedged apart for even monomeric binding to occur.

4.4 Group III Compounds

The symmetrical Group III compounds (DB1746, DB1747, and DB2289) proved to be poor binders of any sequences tested. This suggests that the phenyl-benzimidazole-amidine terminals are much preferred to the phenyl-furan-phenyl-amidine terminals. However, the results from DB1791 indicate that the inclusion of at least one phenyl-furan-phenyl-amidine terminal does have potential as a structural motif. Interestingly, the circular dichroism results (Figure 20) indicate a significant

structural change in the DNA (A_4T_4) upon binding of DB1791. This could be due to the compound folding back on itself, wedging the minor groove apart upon the binding of one or two molecules. Other explanations include significant DNA curvature change, or partial intercalation. If DB1791 is folding back on itself, then the larger, symmetric compounds may also be folding back on themselves. These compounds (DB1746, DB1747, and DB2289) would form larger structures than DB1791 if folded back. The binding sites may not be able to accommodate a larger structure than DB1791, especially considering the large structural changes already visible in the CD of DB1791 and A_4T_4 (Figure 20). The symmetrical compounds, if folded back, would also put the terminal, positively charged amidines closer to each other, which is unfavorable.

If DB1791 either induces a DNA curvature change or partially intercalates, given that only this asymmetric compound binds well out of this group, such a structural change may be feasible only once in the entire binding site. If this change in structure were repeated twice, in close proximity, it could distort the geometry of the DNA to the extent of forming an unstable structure with far from optimal stacking interactions. Such a structure would not be thermodynamically favored, and so a more likely interaction would be only one of the terminals of the symmetric Group III compounds binding. However, if only one of the terminals of DB1746, DB1747, or DB2289 were bound, this would also result in an unstable complex, and the compound would be thermodynamically driven away from minor groove binding.

4.5 Group IV Compounds

Most of the compounds in Group IV also proved to be poor two-site binders. Previously, the replacement of a furan by a thiophene in DB818 was reported to dramatically increase minor groove binding.[3c] However, the inclusion of two such large atoms as sulfur in the ring systems may present a steric obstacle to minor groove binding, specific to larger compounds. Another possibility is that the curvature of the large Group IV compounds is not a good match for the curvature of the minor groove. DB2224 actually destabilized duplex DNA, implying that this compound binds to single stranded DNA better than to double stranded DNA. The cyanine dyes that were selected do not contain any hydrogen bond donors, which is likely one reason they were only able to bind weakly to the tested sequences. The selected cyanine dyes also contain N-CH_3 groups, which may act as steric hindrances to deep minor groove binding. Two of the compounds, DB1255 and DB1998, show some promise as minor groove binders, and may even be able to accommodate two GC base pairs in the two-site model. However, these compounds are also small, and lack the sequence specificity required.

5 CONCLUSIONS

Significantly more information about two-site minor groove binders has been discovered through this work. Since Group I compounds appear to be the strongest binders of A_4T_4, A_4GT_4, and A_4CT_4, slight rigidity or aromaticity in the linker may promote binding. Group II compounds also show good potential, although in general the flexible linker may cause specificity to be more challenging to obtain. The phenyl-

benzimidazole-amidine terminal is another characteristic indicative of increased binding affinity, as well as non-polar amidine extensions.

New ideas have also been developed for compound selectivity and GC recognition. Additional compounds should be developed with conservation of the benzimidazole-amidine terminal, substituting the phenyls for smaller, more sequence-specific groups. In order to distinguish between A_4T_4 and A_4GT_4 or A_4CT_4, more linkers should be designed capable of accepting a hydrogen bond from the protruding amine group of guanine. Replacement of the phenyl moieties by smaller groups may allow such an acceptor to become accessible, and successful hydrogen bond formation could provide a basis for selectivity of A_4GCT_4, A_4GT_4 or A_4CT_4 over A_4T_4. The design of a compound able to bind A_4GCT_4 should include these same characteristics; however, such a compound must have a longer linker and be more able to accommodate the widening of the minor groove and decrease in electrostatic potential such a sequence creates. Additionally, there is a greater chance for compounds to fold back on themselves when the AT sites have more separation. Based on this study, jumping two GC BPs has already proven to be more difficult than jumping one.

Progress for GC selectivity has already been made. DB2120 is selective for sequences containing one intervening GC BP, and this discovery marks a significant step in the development of a new, programmable molecular language for sequence specific minor groove binding compounds. Additionally, replacement of phenyls by pyridines in DB2260 hints of increased GC selectivity, but this should be explored more to determine if such a substitution promotes GC binding with other linking moieties.

Challenges have included limited selectivity among the two-site compounds, and difficulty in studying these large molecules via SPR. There is also a lack of structural information to discern any of the specific binding interactions of these compounds. NMR or crystallographic studies would open the door to more intelligent compound engineering, and a complex with DB2120 would be a good candidate for such work.

A variety of compounds have now been found which bind with a two-site mode. The next step is the development of better linkers capable of GC recognition, and DB2120 is a good lead compound for this work. By incorporating other structural motifs that were found to increase binding, (*i.e.* terminal isopropyl groups, terminal cyclopentane groups, substitution of phenyls by pyridines), stronger binding analogs with this capability are becoming a tangible goal.

REFERENCES

1. (a) Doss, R. M.; Marques, M. A.; Foister, S.; Chenoweth, D. M.; Dervan, P. B., Programmable Oligomers for Minor Groove DNA Recognition. *Journal of the American Chemical Society* **2006,** *128*, 9074-9079; (b) Liu, Y.; Jaiswal, A.; Poggi, M.; Wilson, W. D., Surface Plasmon Resonance and Quartz Crystal Microbalance Methods for Detection of Molecular Interactions. In *Chemosensors: Principles, Strategies, and Applications*, 1 ed.; Wang, B.; Anslyn, E. V., Eds. John Wiley & Sons, Inc.: 2011; pp 329-343.

2. (a) Dabrowiak, J. C., *Metals in Medicine*. Wiley: West Sussex, 2009; (b) Nguyen, B.; Tanious, F. A.; Wilson, W. D., Biosensor-surface Plasmon Resonance: Quantitative Analysis of Small Molecule-nucleic Acid Interactions. *Methods* **2007,** *42*, 150-161.

3. (a) Rahimian, M.; Kumar, A.; Say, M.; Bakunov, S. A.; Boykin, D. W.; Tidwell, R. R.; Wilson, W. D., Minor Groove Binding Compounds That Jump a GC Base Pair and Bind to Adjacent AT Base Pair Sites. *Biochemistry* **2009,** *41*, 1573-1583; (b) Roy Chowdhury, A.; Bakshi, R.; Wang, J.; Yildirir, G.; Liu, B.; Pappas-Brown, V.; Tolun, G.; Griffith, J. D.; Shapiro, T. A.; Jensen, R. E.; Englund, P. T., The Killing of African Trypanosomes by Ethidium Bromide. *PLoS Pathog* **2010,** *6* (12), e1001226; (c) Wilson, W. D.; Nguyen, B.; Tanios, F. A.; Mathis, A.; Hall, J. E.; Stephens, C. E.; Boykin, D. W., Dications That Target the DNA Minor Groove: Compund Design and Preparation, DNA Interactions, Cellular Distribution and Biological Activity. *Current Medicinal Chemistry Anti-Cancer Agents* **2005,** *5*, 389-408.

4. *Nucleic Acids in Chemistry and Biology*. 2 ed.; Oxford University Press: New York, 1996.

5. Sinden, R. R., *DNA Structure and Function*. Academic Press, Inc.: San Diego, CA, 1994.

6. (a) Lindemose, S.; Nielsen, P. E.; Hansen, M.; Møllegaard, N. E., A DNA Minor Groove Electronegative Potential Genome Map Based on Photo-chemical Probing. *Nucleic Acids Research* **2011**; (b) Biship, E. P.; Rohs, M.; Parker, S. C. J.; West, S. M.; Liu, P.; Mann, R. S.; Honig, B.; Tullius, T. D., A Map of Minor Groove Shape and Electrostatic Potential from Hydroxyradical Cleavage Patterns of DNA. *ACS Chemical Biology* **2011,** *6*, 1314-1320.

7. (a) Gregson, S. J.; Howard, P. W.; Gullick, D. R.; Hamaguchi, A.; Corcoran, K. E.; Brooks, N. A.; Hartley, J. A.; Jenkins, T. C.; Patel, S.; Guille, M. J.; Thurston, D. E., Linker Length Modulates DNA Cross-Linking Reactivity and Cytotoxic Potency of C8/C8' Ether-Linked C*2*-*exo*-Unsaturated Pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD) Dimers. *Journal of Medicinal Chemistry* **2004,** *47*, 1161-1174; (b) Rosado, H.; Rahman, K. M.; Feuerbaum, E.-A.; Hinds, J.; Thurston, D. E.; Taylor, P. W., The Minor Groove-binding Agent ELB-21 Forms Multiple Interstrand and Intrastrand Covalent Cross-links with Duplex DNA and Displays Potent Bactericidal Activity Against Methicillin-resistant *Staphylococcus aureus*. *Journal of Antimicrobial Chemotherapy* **2011**.

8. Tanada, M.; Sasaki, S., Discrimination of the Length of Two Remote Binding Sites by the Spacer-linked DNA Minor Groove Binders. *Nucleic Acids Symposium Series No. 49* **2005**, 1-2.

9. (a) Hawekins, C. A.; Baird, E. E.; Dervan, P. B.; Wemmer, D. E., Analysis of Hairpin Polyamide Complexes Having DNA Binding Sites in Close Proximity. *Journal of the American Chemical Society* **2002,** *124*, 12689-12696; (b) Marques, M. A.; Doss, R. M.; Foister, S.; Dervan, P. B., Expaning the Repertoire of Heterocycle Ring Pairs for Programmable Minor Groove DNA Recognition. *Journal of the American Chemical Society* **2004,** *126*, 10339-10349; (c) Weyermann, P.; Dervan, P. B., Recognition of Ten Base Pairs of DNA by Head-to-Head Hairpin Dimers. *Journal of the American Chemical Society* **2002,** *124*, 6872-6878.

10. (a) Trauger, J. W.; Baird, E. E.; Dervan, P. B., Recognition of 16 Base Pairs in the Minor Groove of DNA by a Pyrrole-Imidazole Polyamide Dimer. *Journal of the American Chemical Society* **1998,** *120* (3534-3535); (b) Trauger, J. W.; Baird, E. E.;

Mrksich, M.; Dervan, P. B., Extension of Sequence-Specific Recognition in the Minor Groove of DNA by Pyrrole-Imidazole Polyamids to 9-13 Base Pairs. *Journal of the American Chemical Society* **1996,** *118*, 6160-6166.

11. (a) Cordonnier, M.-H. D.; Hilderbrand, M.-P.; Baldeyrou, B.; Lansiaux, A.; Keuser, C.; Benzshawel, K.; Lemster, T.; Pindur, U., Deisgn, Synthesis and Biological Evaluation of New Oligopyrrole Carboxyamides Linked with Tricyclic DNA-intercalators as Potential DNA Ligand or Tomoisomerase Inhibitors. *European Journal of Medicinal Chemistry* **2007,** *42*, 752-771; (b) Zhang, R.; Wu, X.; Guziec, L. J.; Guziec, F. S.; Chee, G.-L.; Yalowich, J. C.; Hasinoff, B. B., Design, Synthesis and Biological Evaluation of a Novel Series of Anthapyrazoles Linked with Netropsin-like Oligopyrrole Carboxamides as Anticancer Agents. *Bioorganic & Medicinal Chemistry* **2010,** *18*, 3974-3984; (c) Waring, M.; Bailly, C., DNA Recognition by Intercalators and Hybrid Molecules. *Journal of Molecular Recognition* **1994,** *7*, 109-122.

12. Mirkin, S., Expandible DNA Repeats and Human Disease. *Nature* **2007,** *447*, 932-940.

13. Wilson, W. D.; Tanious, T. A.; Fernandez-Saiz, M.; Rigl, C. T., Evaluation of Drug-Nucleic Acid Interactions by Thermal Melting Curves. In *Methods in Molecular Biology Vol. 90: Drug-DNA Interaction Protocols*, Fox, K. R., Ed. Humana Press Inc.: Totowa, 1997.

14. (a) Tanious, F. A.; Nguyen, B.; Wilson, W. D., Biosensor-Surface Plasmon Resonance Methods for Quantitative Analysis of Biomolecular Interactions. In *Methods in Cell Biology*, Elsevier Inc.: 2008; Vol. 84, pp 53-77; (b) Wilson, W. D., Analyzing Biomolecular Interactions. *Science* **2002,** *295*, 2103-2105.

15. *Spartan '04 Wilndows Tutorial and User's Guide*. Wavefunction, Inc.: Irvine, 2004.

16. Neslon, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A., The Structure of an Oligo(dA)-oligo(dT) Tract and its Biological Implications. *Nature* **1987,** *330*, 221-226.

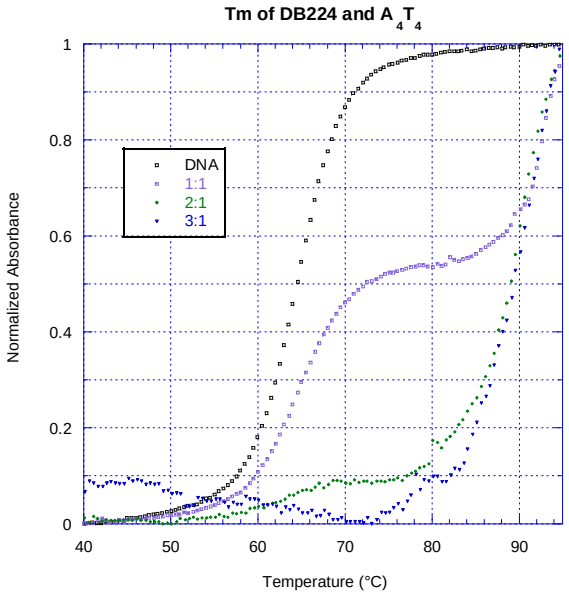
17. Bansal, M., DNA Structure: Revisiting the Watson-Crick Double Helix. *Current Science* **2003,** *85* (11), 1556-1563.

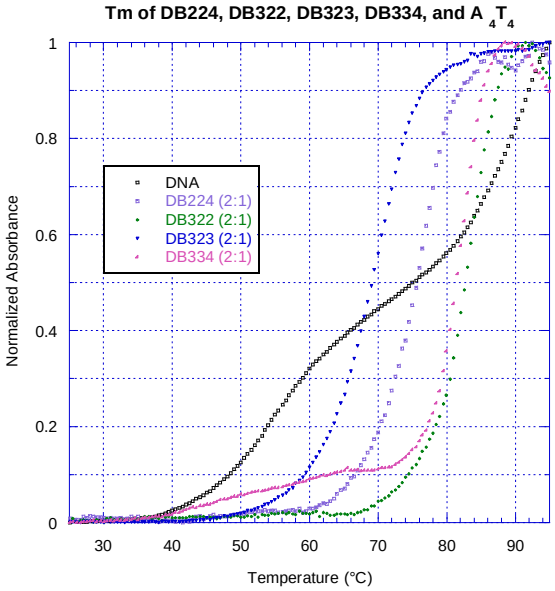
18. Liu, Y.; Chai, Y.; Kumar, A.; Boykin, D.; Wilson, W., Designed Compounds for Recognition of 10 Base Pairs of DNA with Two AT Binding Sites. *Journal of the American Chemical Society* **2012,** *134* (11), 5290-9.

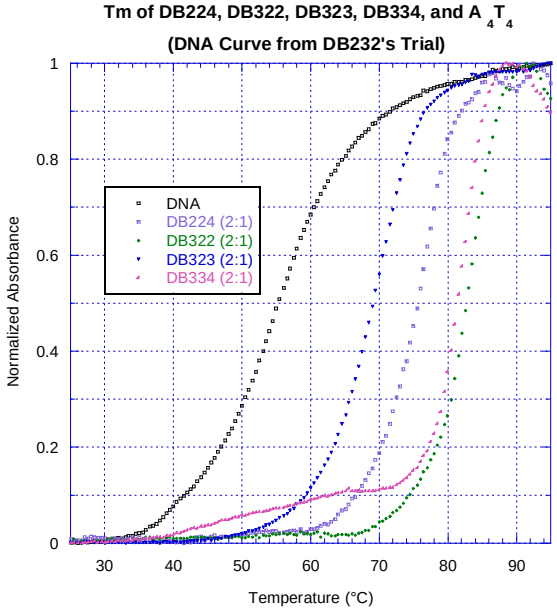
APPENDIX

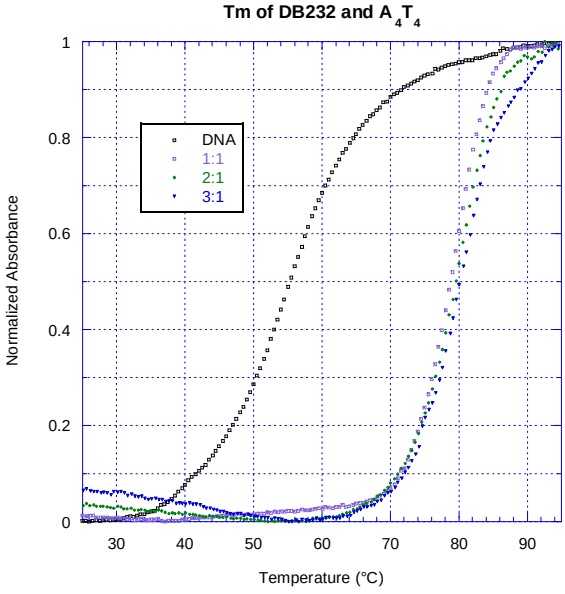
The change in thermal melting temperature was used as a screening method for binding affinity. Typically, a ratio of 2:1 was used for comparison. In order to speed up data collection, several compounds were often tested at a single ratio with DNA. For each experiment, a graph of normalized absorbance vs. temperature was created. A combination of the derivative function and estimation from the graphs was used to determine the ΔTm values listed in Table 1. Initial DNA absorbance was typically 0.6-0.8, and final DNA absorbance was typically 0.8-1.0. Hyperchromicity was approximately 24%.

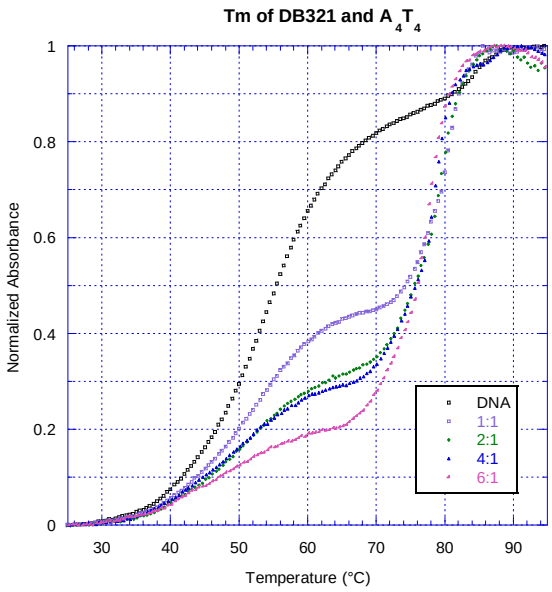
A_4T_4:

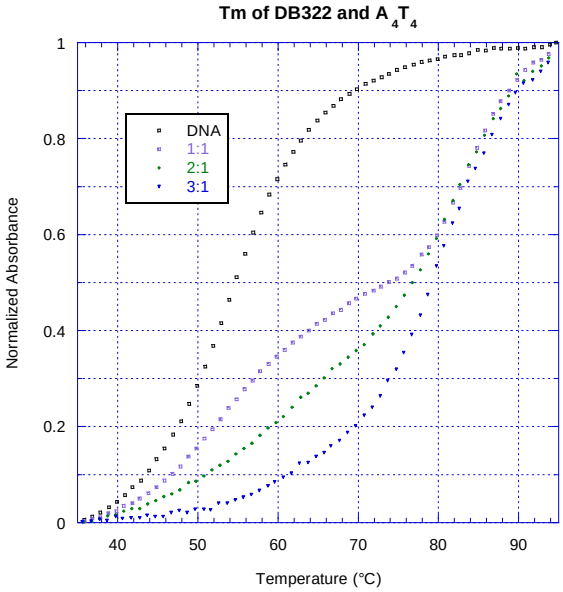

Tm of DB224 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

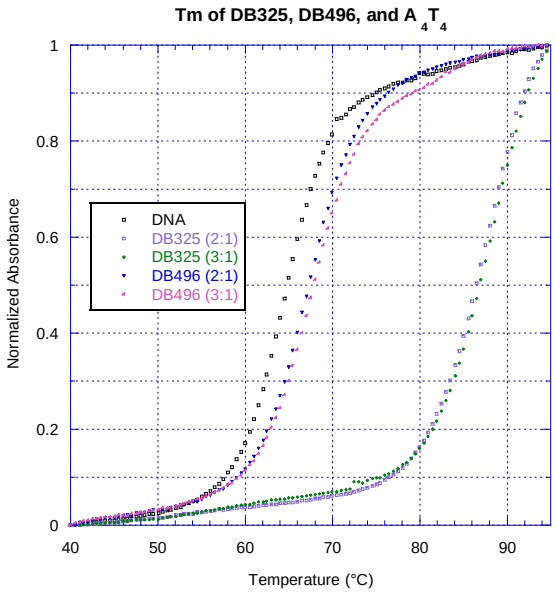

Tm of DB224, DB322, DB323, DB334, and A4T4
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB322 (2:1)
DB323 (2:1)
DB334 (2:1)

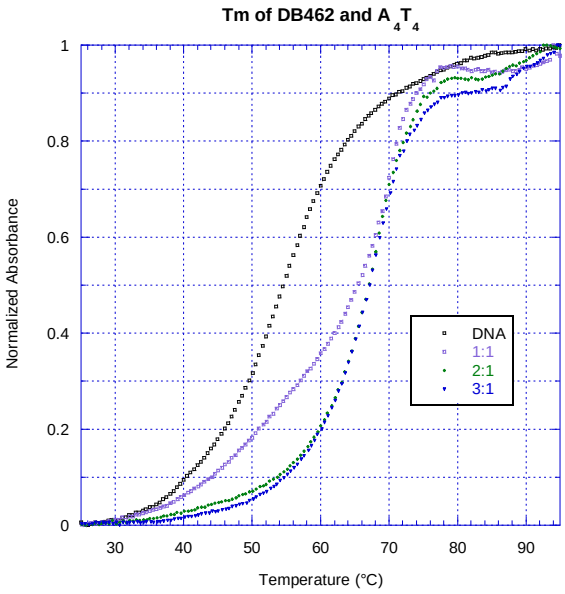

Tm of DB224, DB322, DB323, DB334, and A4T4
(DNA Curve from DB232's Trial)
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB322 (2:1)
DB323 (2:1)
DB334 (2:1)

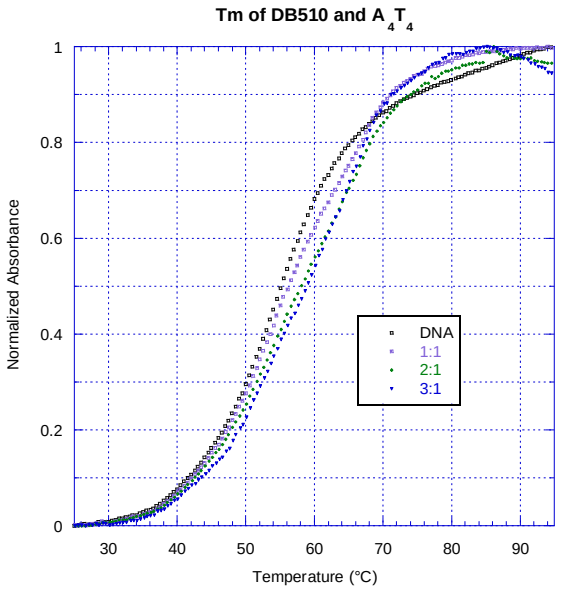

Tm of DB232 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

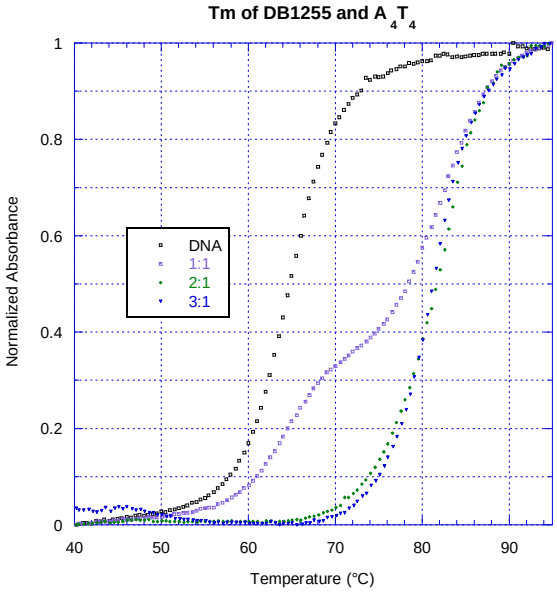

Tm of DB321 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
4:1
6:1

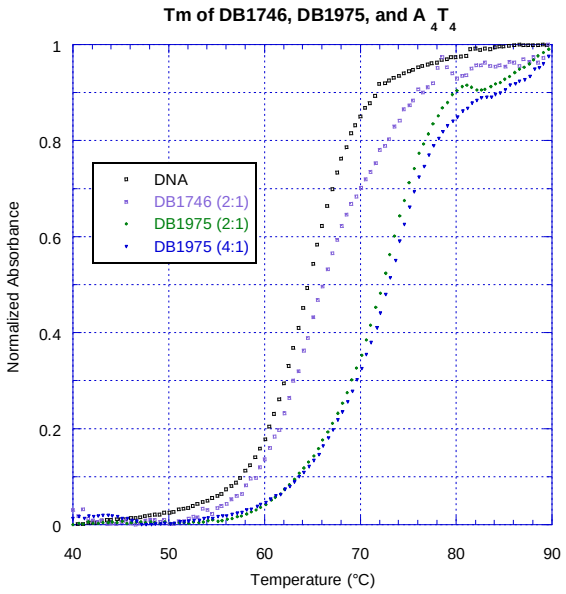

Tm of DB322 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

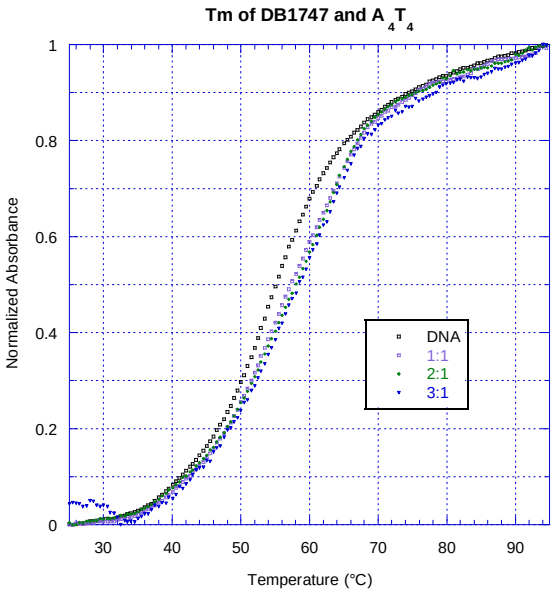

Tm of DB325, DB496, and A4T4
Normalized Absorbance
Temperature (°C)
DNA
DB325 (2:1)
DB325 (3:1)
DB496 (2:1)
DB496 (3:1)

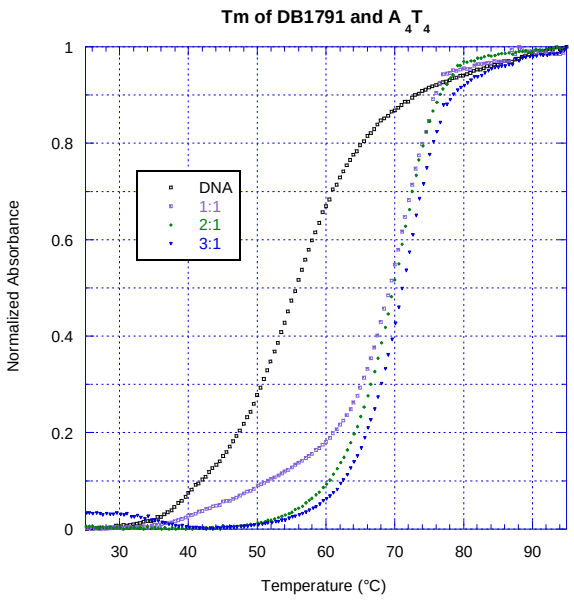

Tm of DB462 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

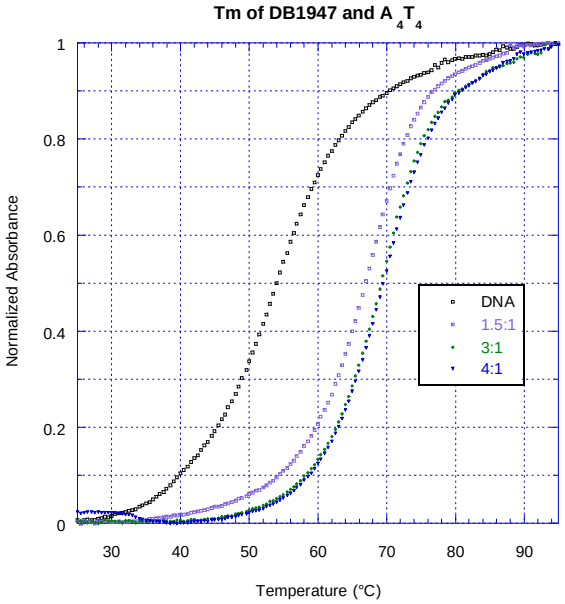

Tm of DB510 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

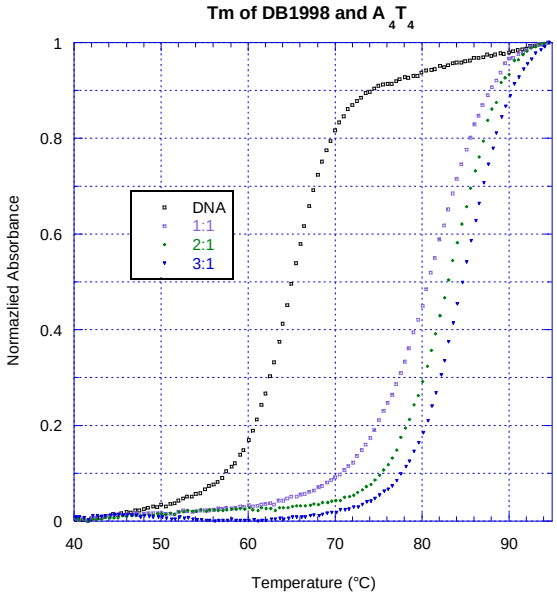

Tm of DB1255 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

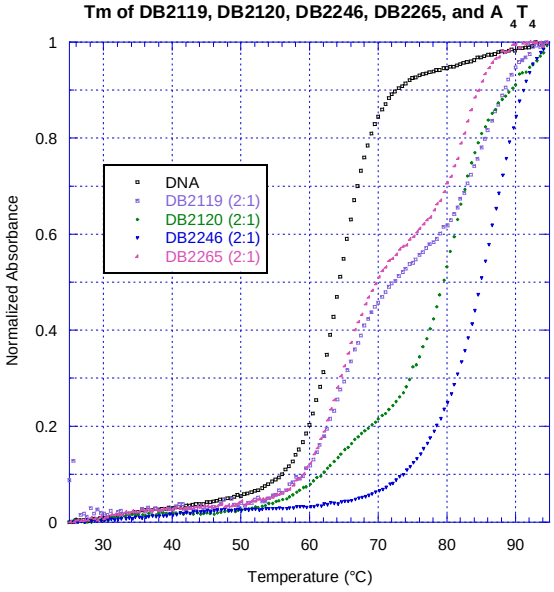

Tm of DB1746, DB1975, and A4T4
Normalized Absorbance
Temperature (°C)
DNA
DB1746 (2:1)
DB1975 (2:1)
DB1975 (4:1)

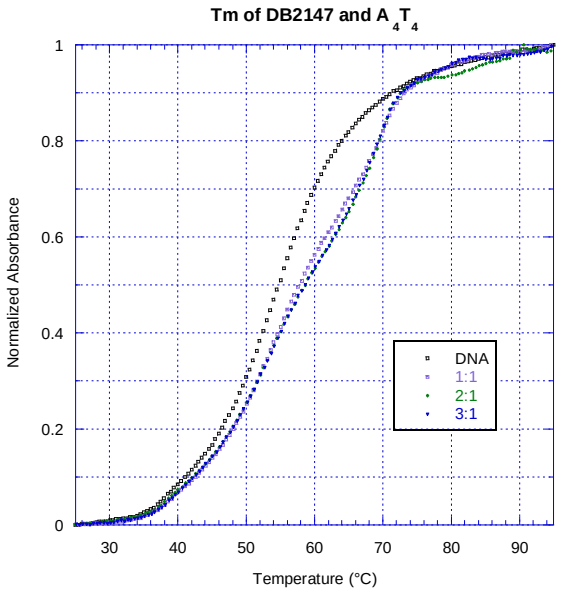

Tm of DB1747 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

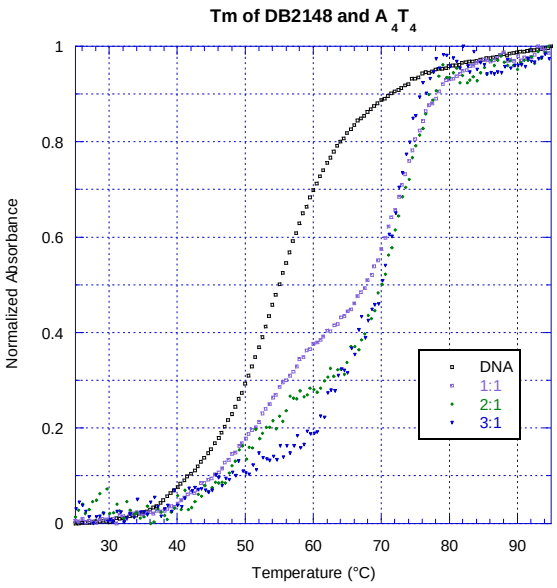

Tm of DB1791 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

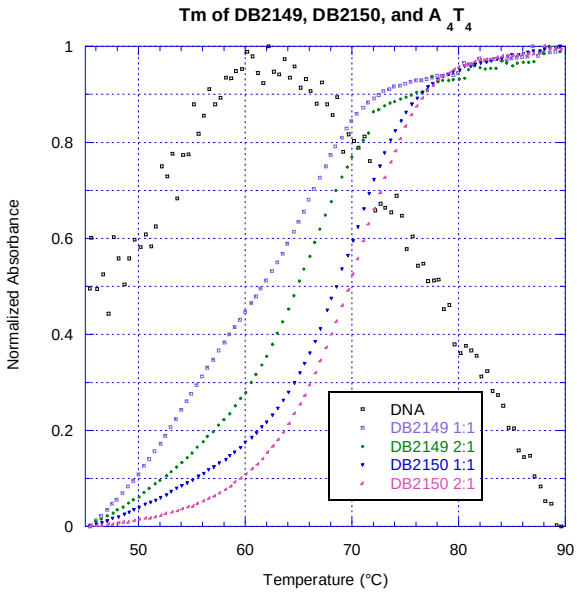

Tm of DB1947 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1.5:1
3:1
4:1

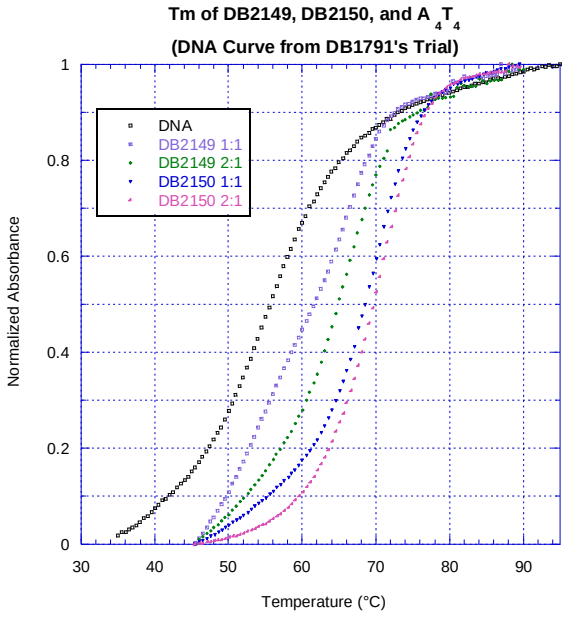

Tm of DB1998 and A4T4
Normazlied Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

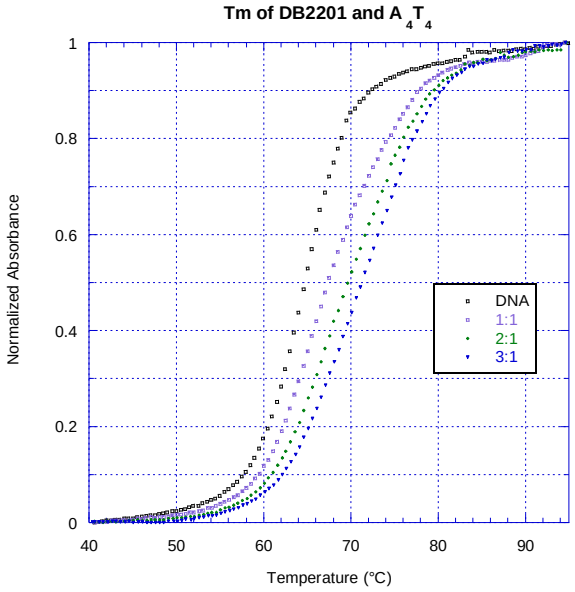

Tm of DB2119, DB2120, DB2246, DB2265, and A4T4
Normalized Absorbance
Temperature (°C)
DNA
DB2119 (2:1)
DB2120 (2:1)
DB2246 (2:1)
DB2265 (2:1)

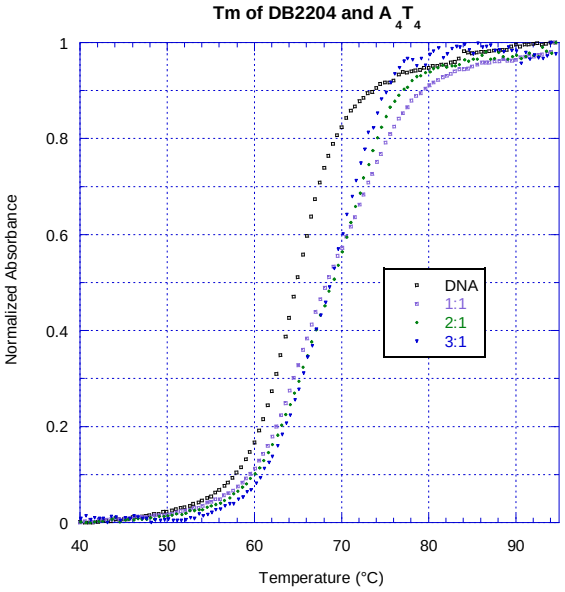

Tm of DB2147 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

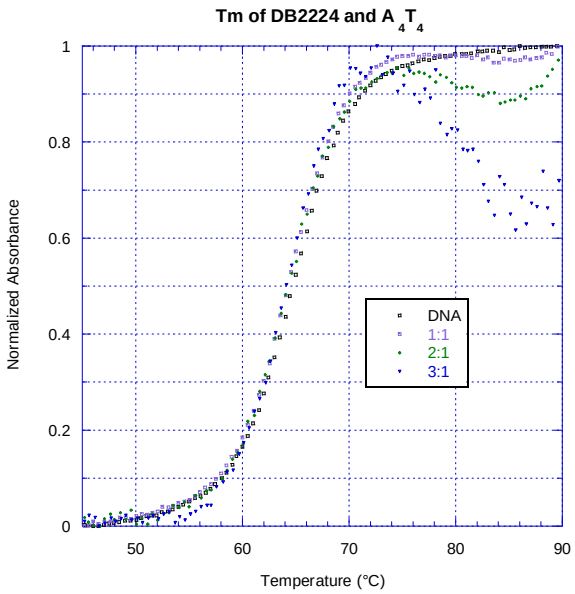

Tm of DB2148 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

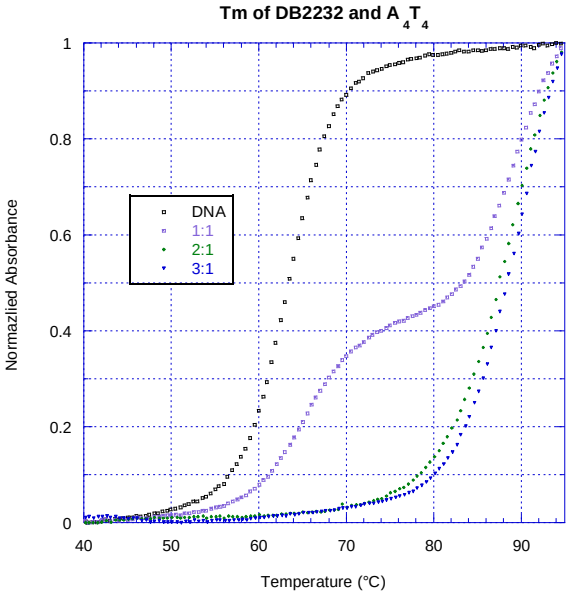

Tm of DB2149, DB2150, and A4T4
Normalized Absorbance
Temperature (°C)
DNA
DB2149 1:1
DB2149 2:1
DB2150 1:1
DB2150 2:1

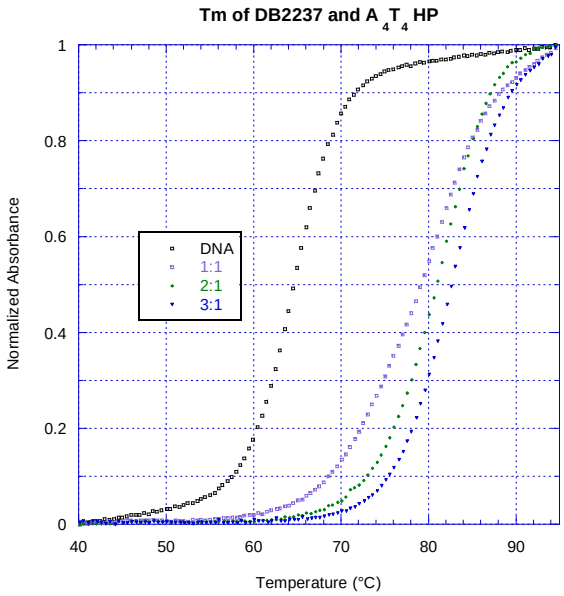

Tm of DB2149, DB2150, and A4T4
(DNA Curve from DB1791's Trial)
Normalized Absorbance
Temperature (°C)
DNA
DB2149 1:1
DB2149 2:1
DB2150 1:1
DB2150 2:1

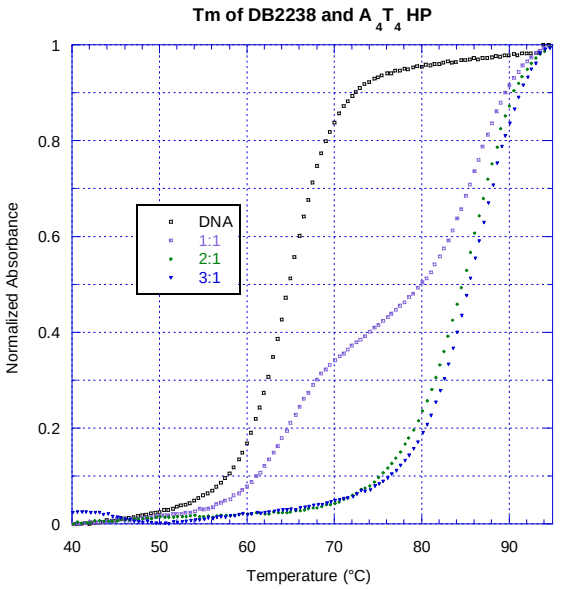

Tm of DB2201 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

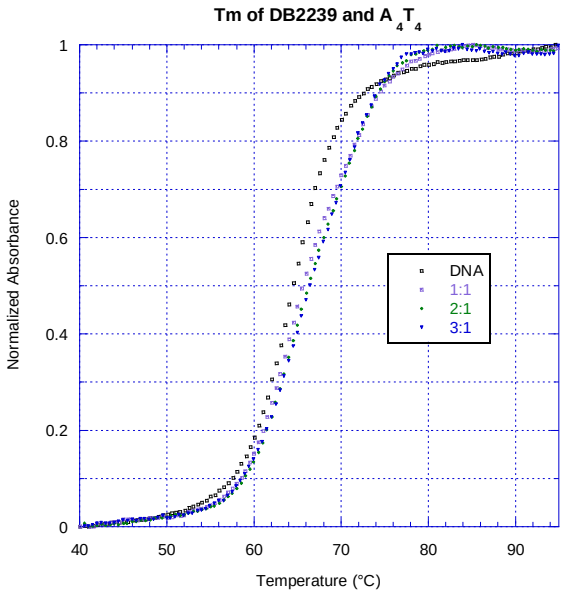

Tm of DB2204 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

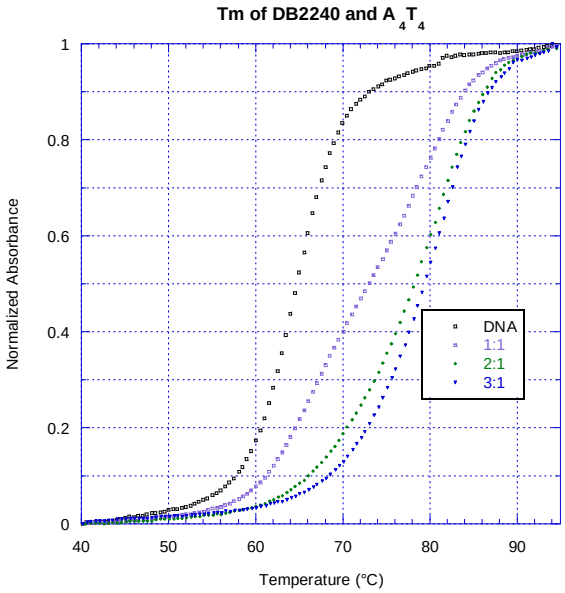

Tm of DB2224 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

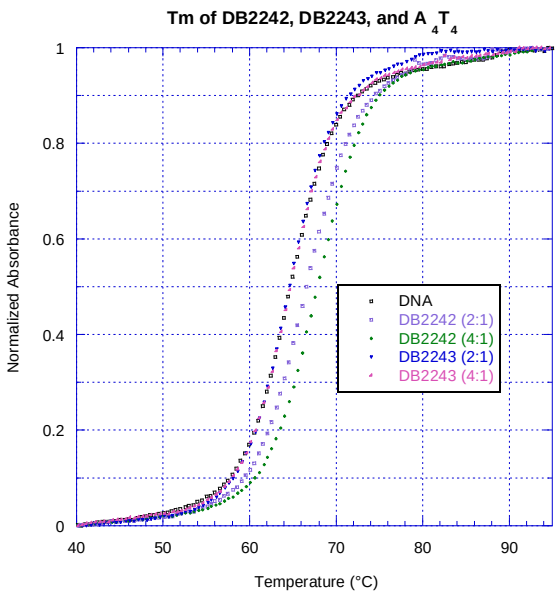

Tm of DB2232 and A4T4
Normazlied Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

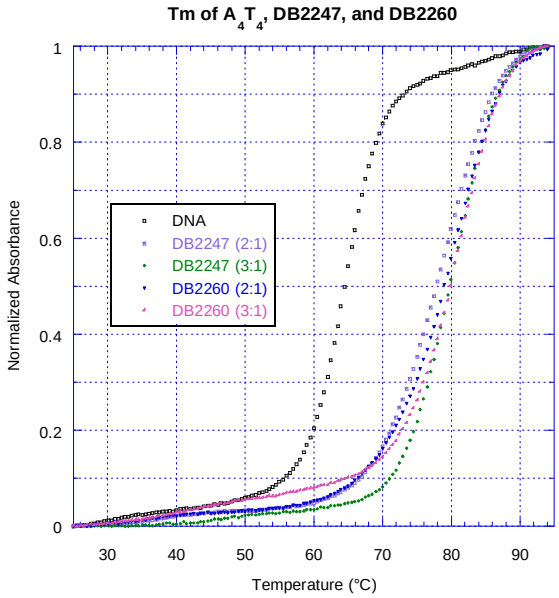

Tm of DB2237 and A4T4 HP
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

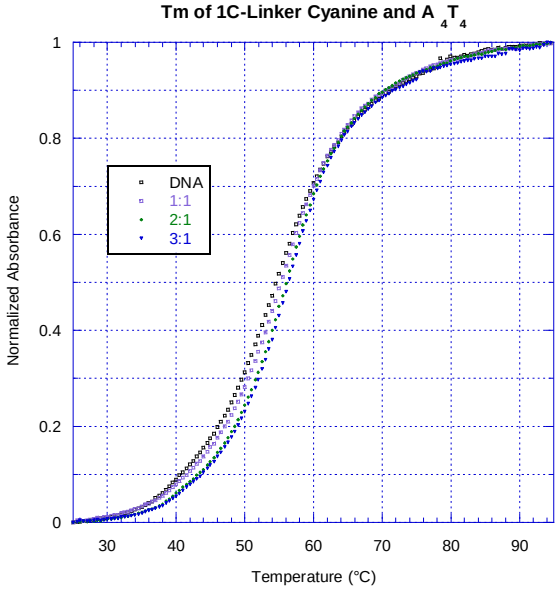

Tm of DB2238 and A4T4 HP
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

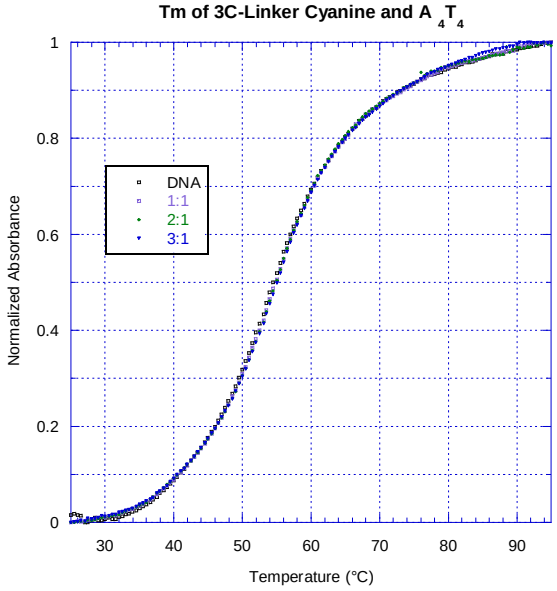

Tm of DB2239 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

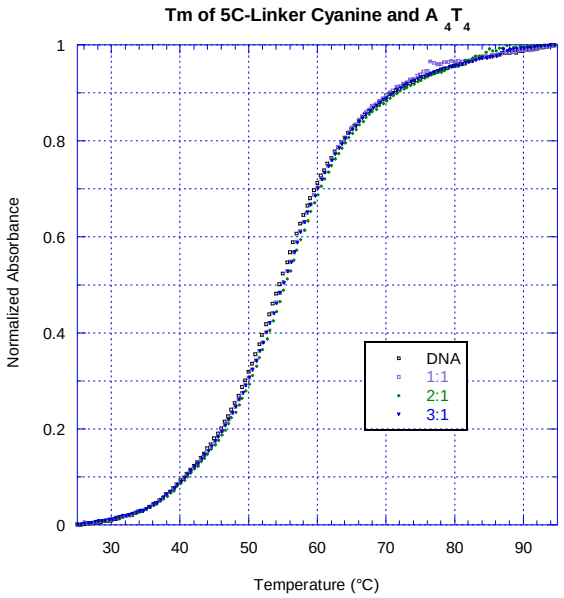

Tm of DB2240 and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1


Tm of DB2242, DB2243, and A4T4
Normalized Absorbance
Temperature (°C)
DNA
DB2242 (2:1)
DB2242 (4:1)
DB2243 (2:1)
DB2243 (4:1)

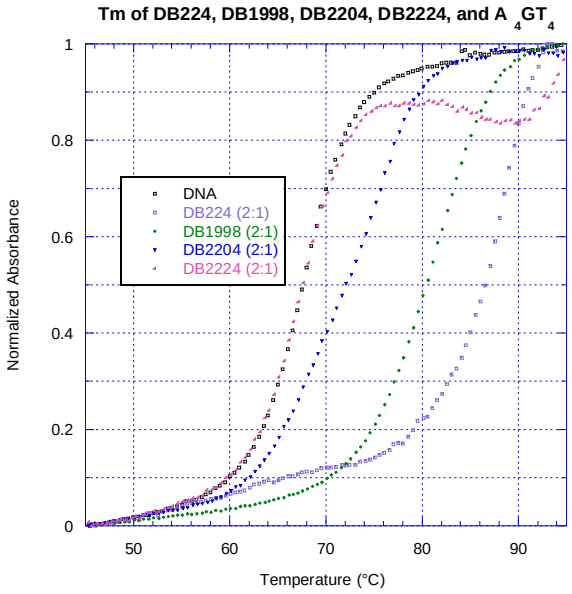

Tm of A4T4, DB2247, and DB2260
Normalized Absorbance
Temperature (°C)
DNA
DB2247 (2:1)
DB2247 (3:1)
DB2260 (2:1)
DB2260 (3:1)

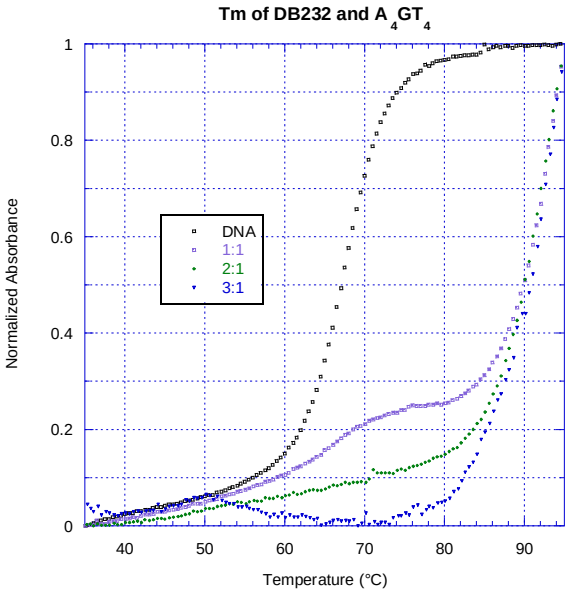

Tm of 1C-Linker Cyanine and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

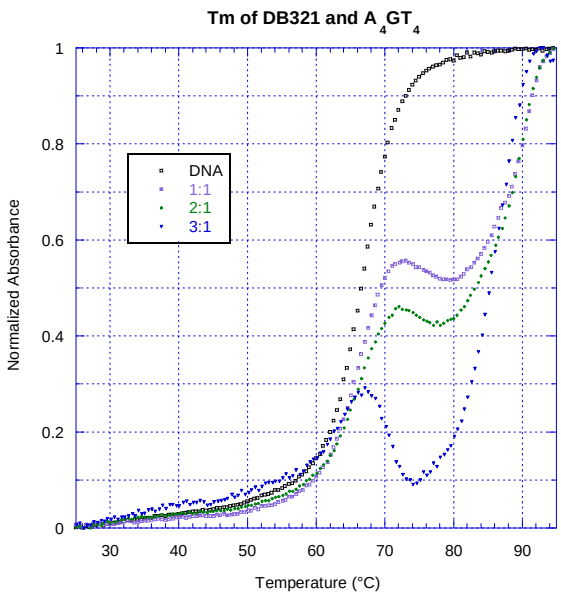

Tm of 3C-Linker Cyanine and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

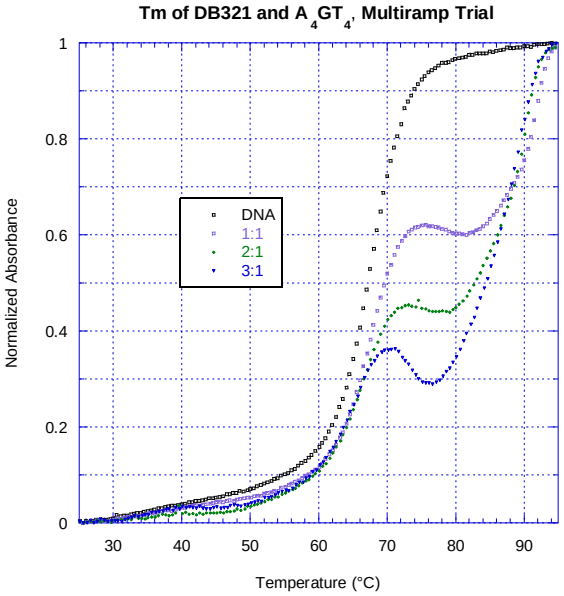

Tm of 5C-Linker Cyanine and A4T4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

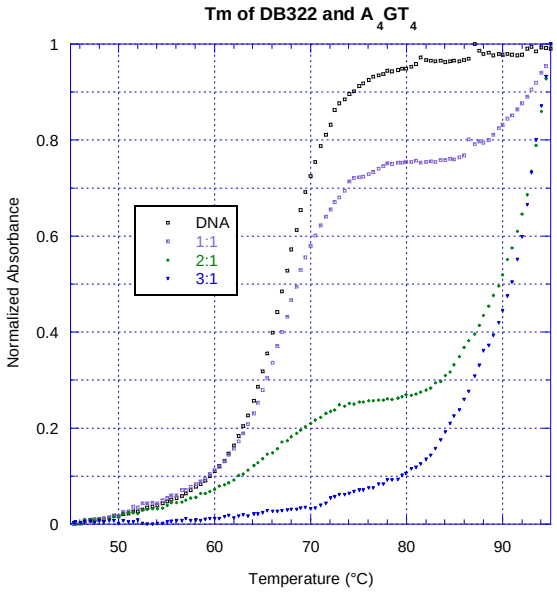
A_4GT_4:

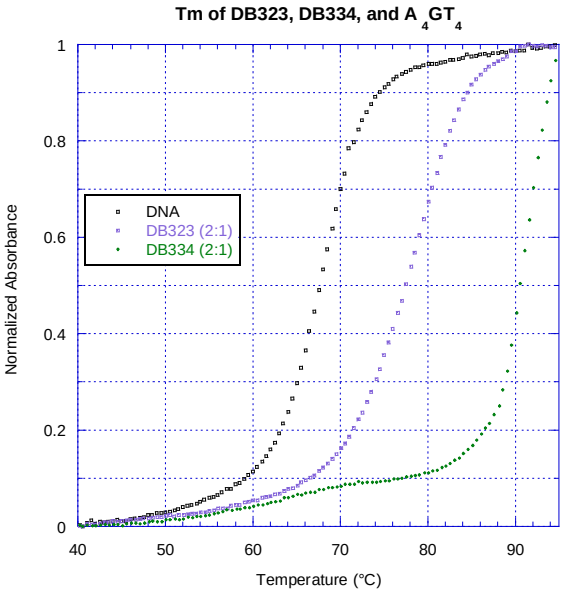

Tm of DB224, DB1998, DB2204, DB2224, and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB1998 (2:1)
DB2204 (2:1)
DB2224 (2:1)

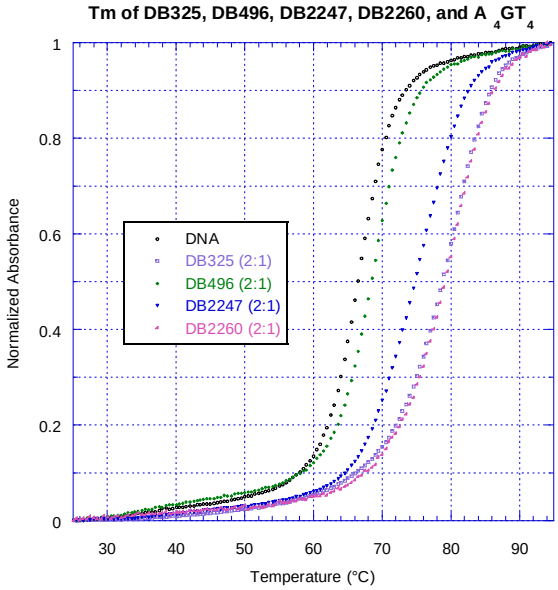

Tm of DB232 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

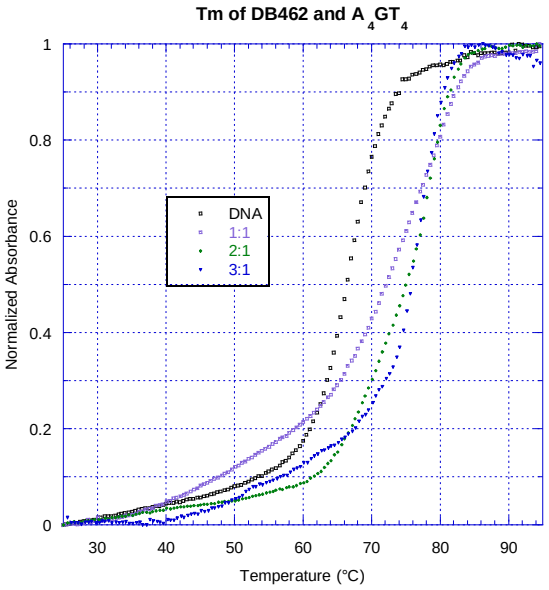

Tm of DB321 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

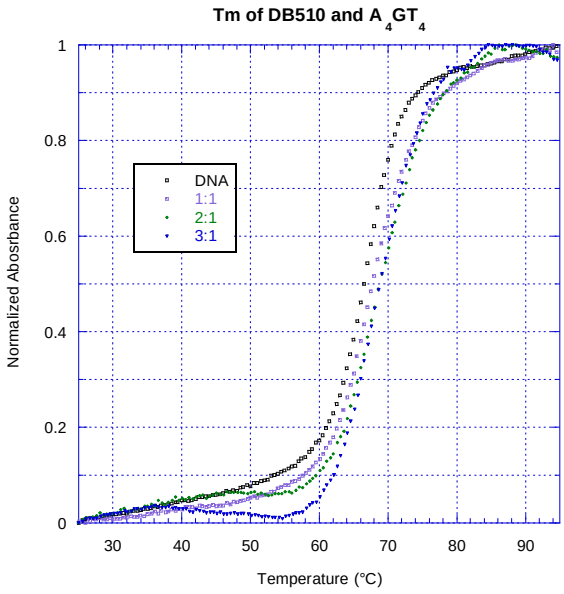

Tm of DB321 and A4GT4, Multiramp Trial
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

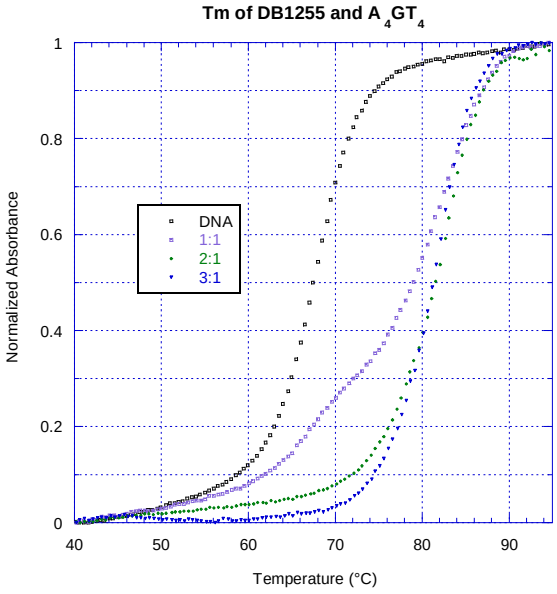

Tm of DB322 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

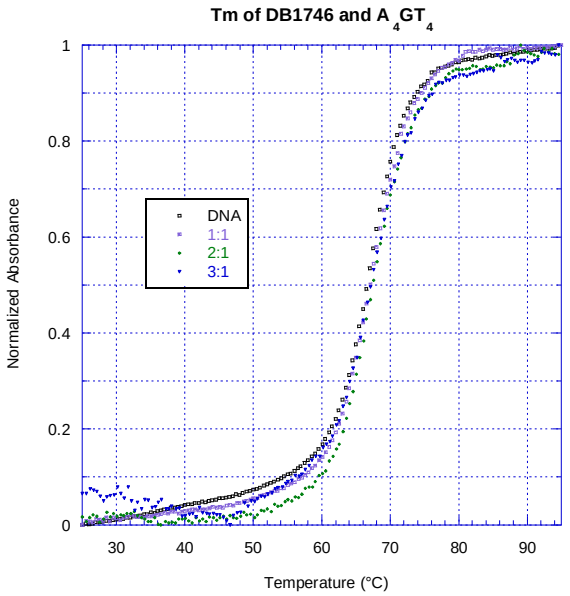

Tm of DB323, DB334, and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
DB323 (2:1)
DB334 (2:1)

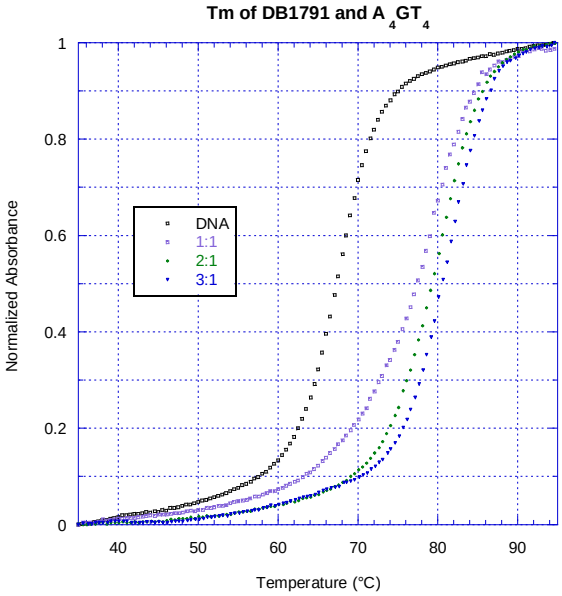

Tm of DB325, DB496, DB2247, DB2260, and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
DB325 (2:1)
DB496 (2:1)
DB2247 (2:1)
DB2260 (2:1)

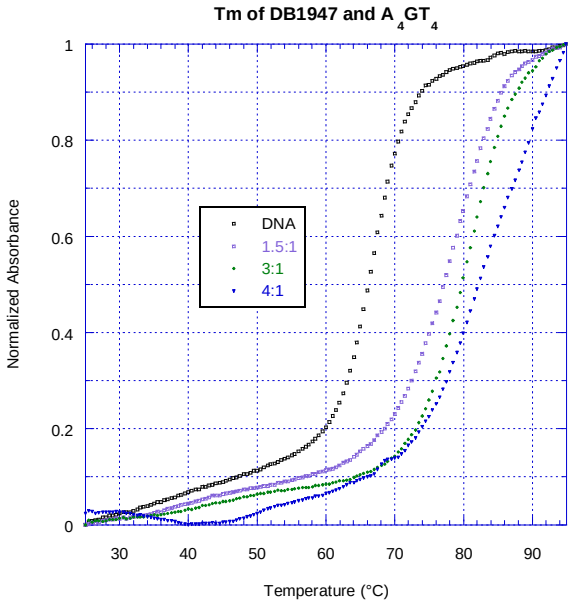

Tm of DB462 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

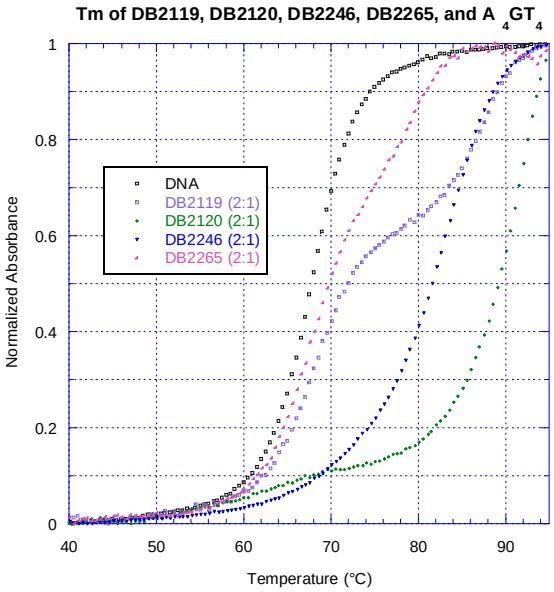

Tm of DB510 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

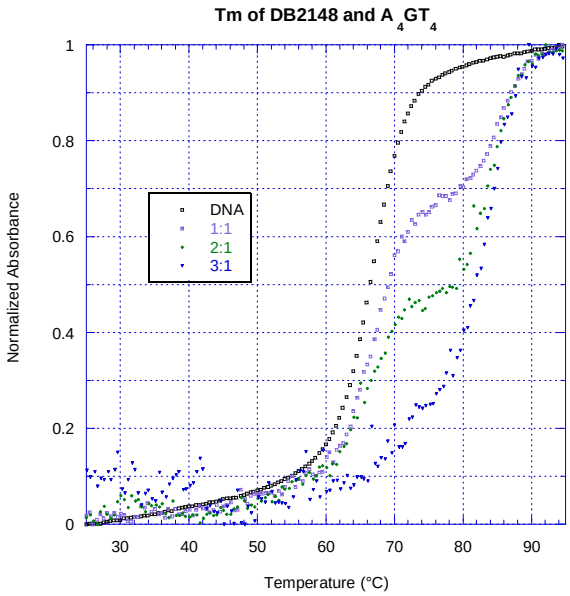

Tm of DB1255 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

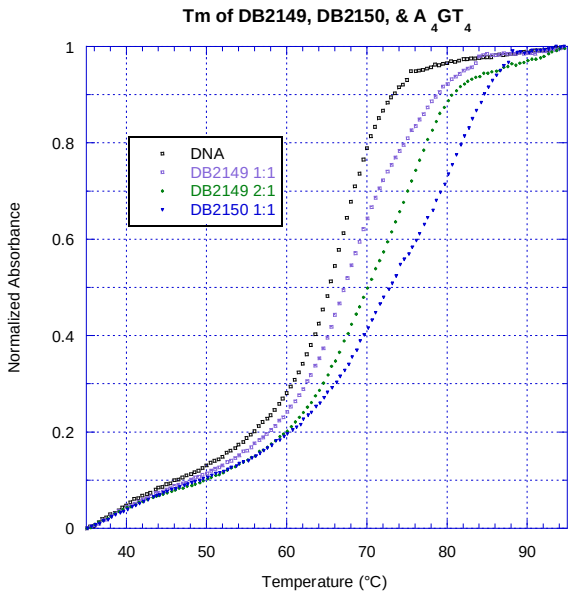

Tm of DB1746 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

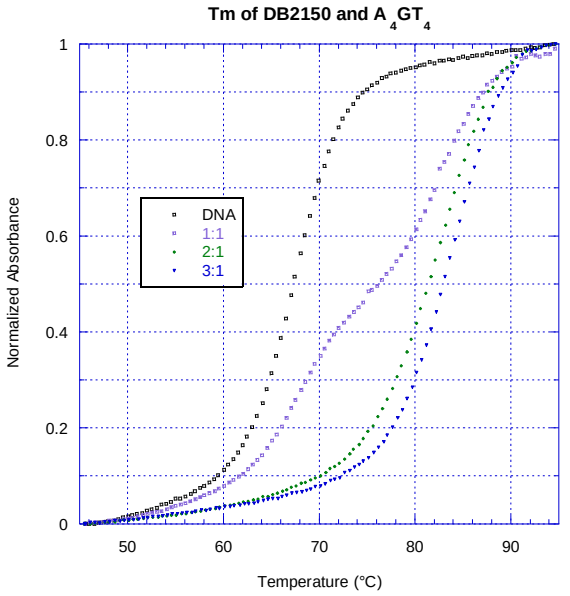

Tm of DB1791 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

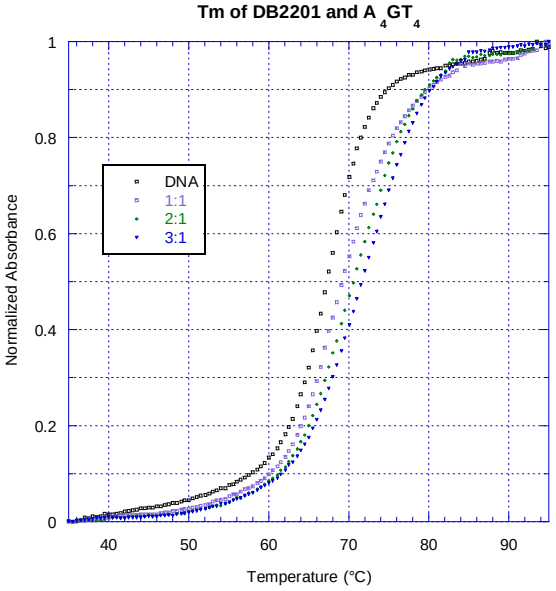

Tm of DB1947 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1.5:1
3:1
4:1

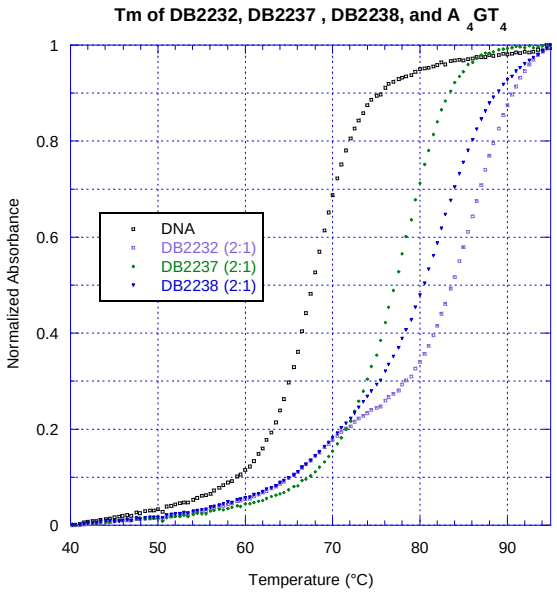

Tm of DB2119, DB2120, DB2246, DB2265, and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
DB2119 (2:1)
DB2120 (2:1)
DB2246 (2:1)
DB2265 (2:1)

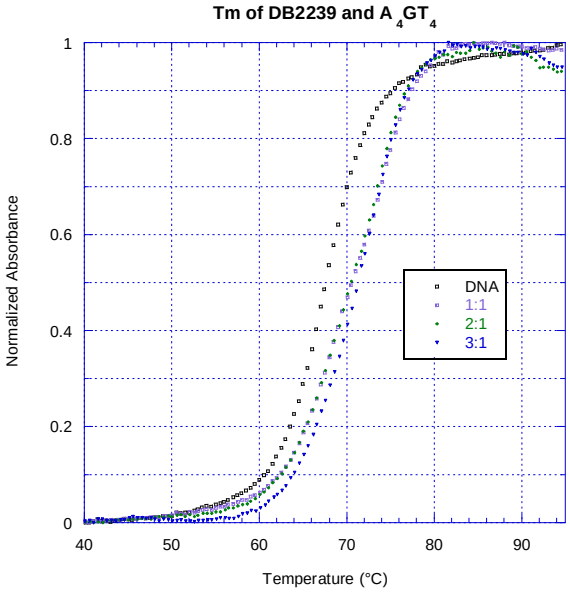

Tm of DB2148 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

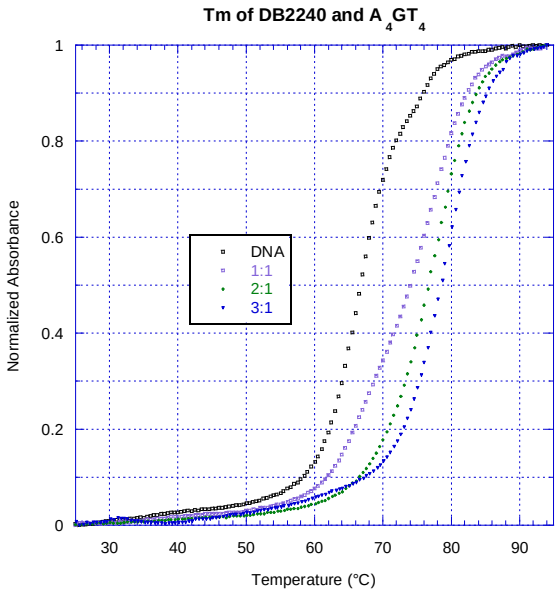

Tm of DB2149, DB2150, & A4GT4
Normalized Absorbance
Temperature (°C)
DNA
DB2149 1:1
DB2149 2:1
DB2150 1:1

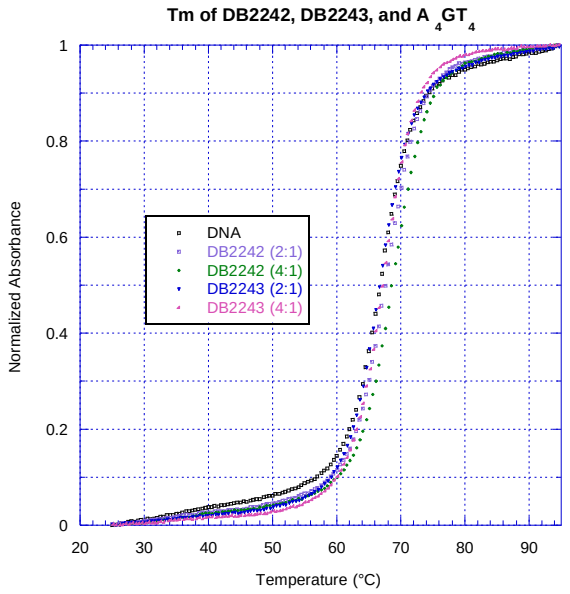

Tm of DB2150 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

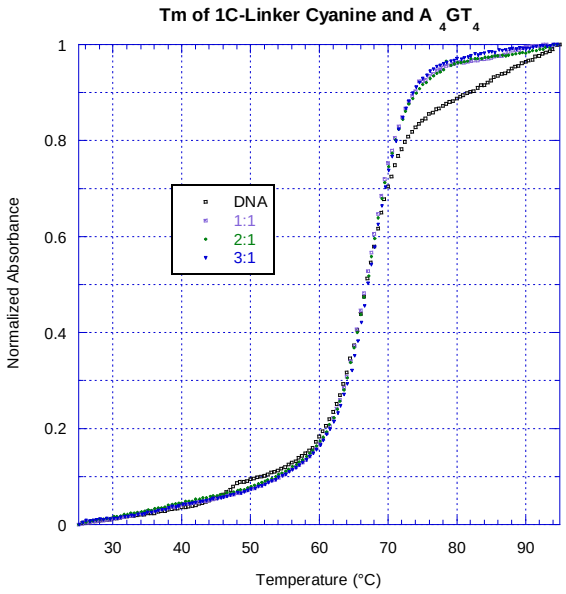

Tm of DB2201 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

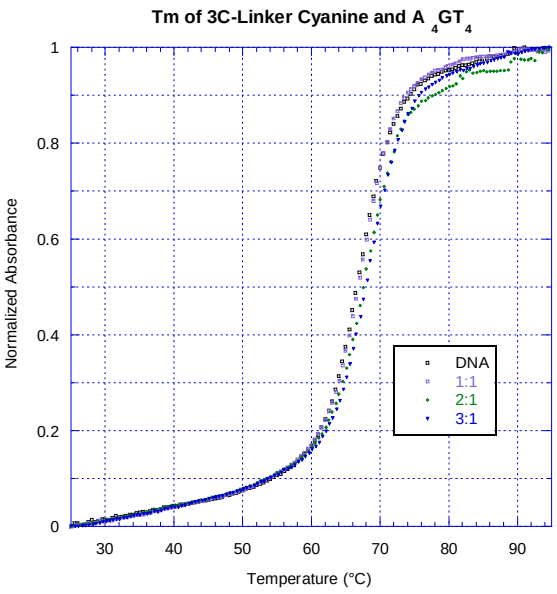

Tm of DB2232, DB2237 , DB2238, and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
DB2232 (2:1)
DB2237 (2:1)
DB2238 (2:1)

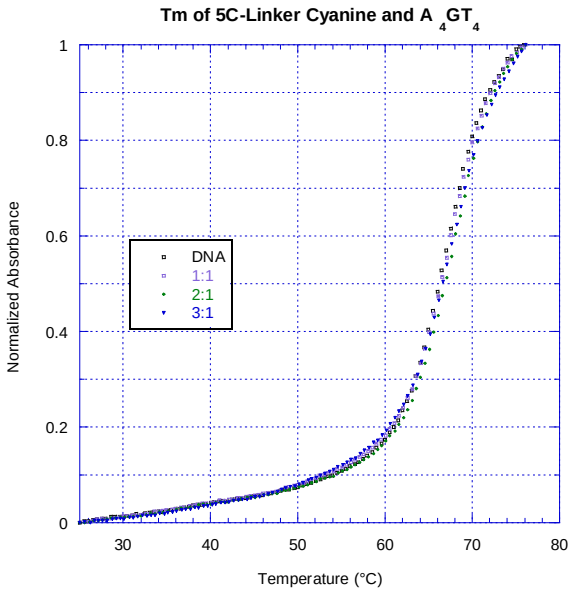

Tm of DB2239 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1


Tm of DB2240 and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

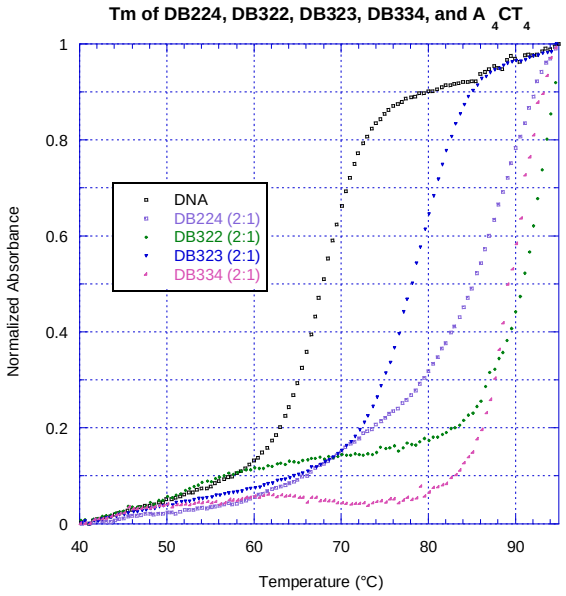

Tm of DB2242, DB2243, and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
DB2242 (2:1)
DB2242 (4:1)
DB2243 (2:1)
DB2243 (4:1)

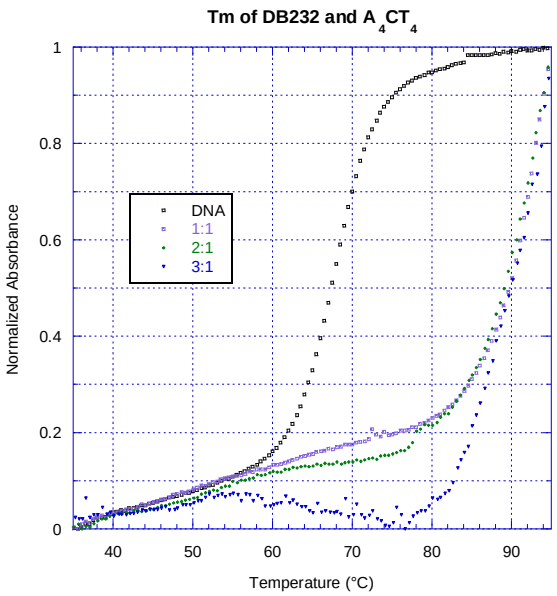

Tm of 1C-Linker Cyanine and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1
0
0.2
0.4
0.6
0.8
1
30
40
50
60
70
80
90

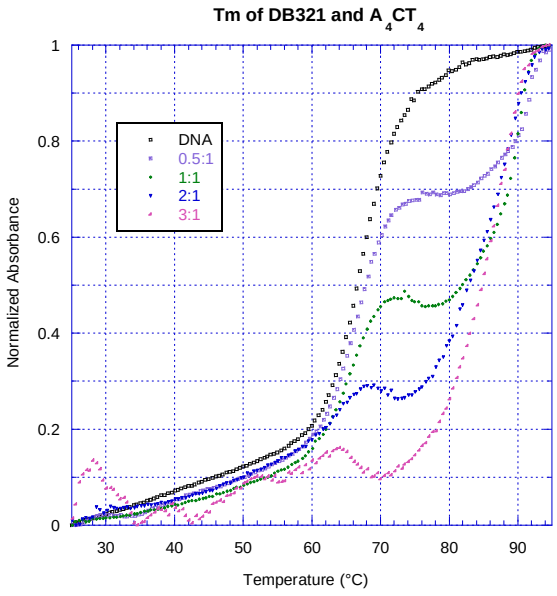

Tm of 3C-Linker Cyanine and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1
0
0.2
0.4
0.6
0.8
1
30
40
50
60
70
80
90

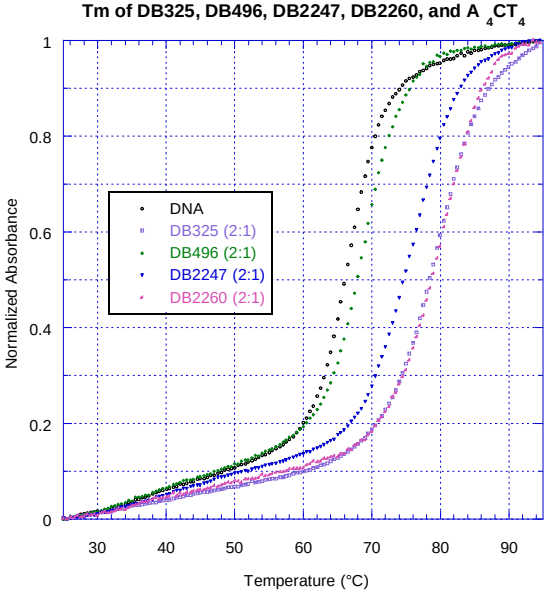

Tm of 5C-Linker Cyanine and A4GT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1
0
0.2
0.4
0.6
0.8
1
30
40
50
60
70
80

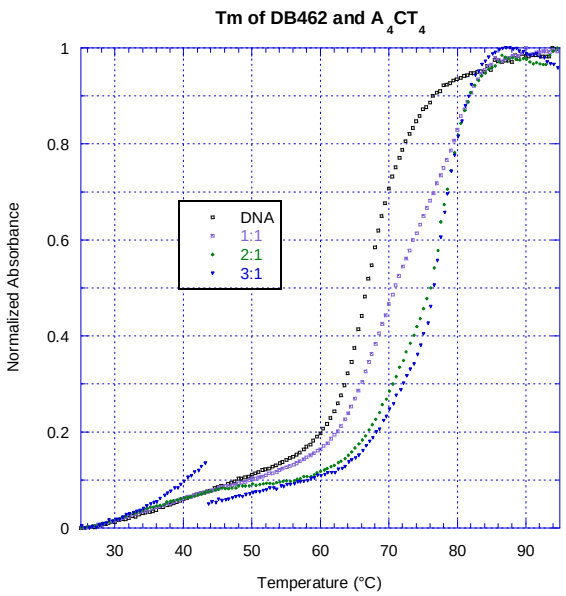
A_4CT_4:

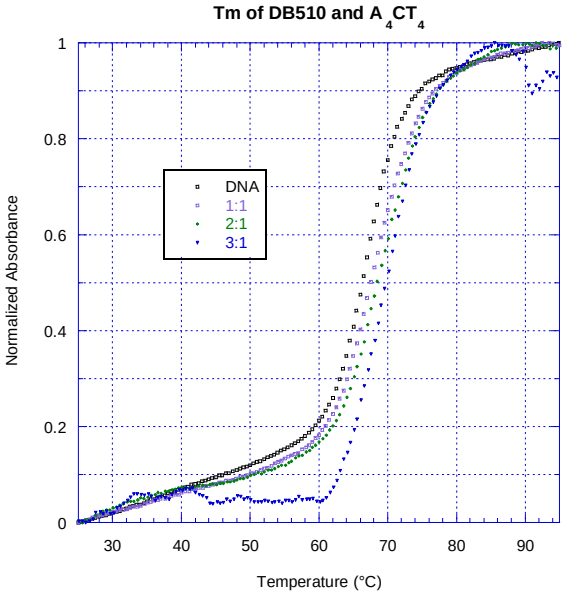

Tm of DB224, DB322, DB323, DB334, and A 4CT4
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB322 (2:1)
DB323 (2:1)
DB334 (2:1)

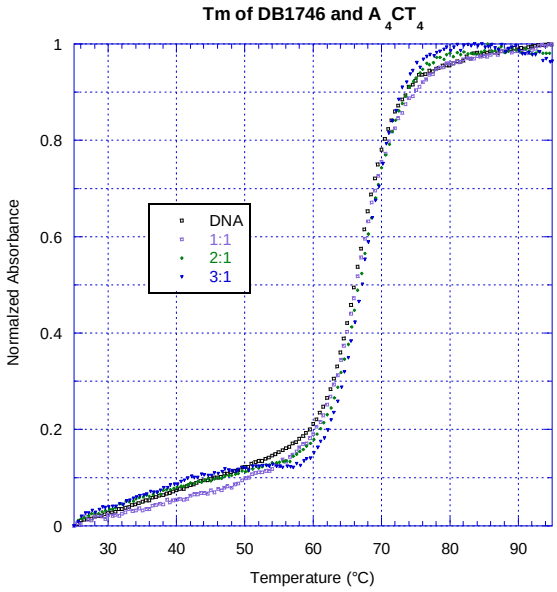

Tm of DB232 and A 4CT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

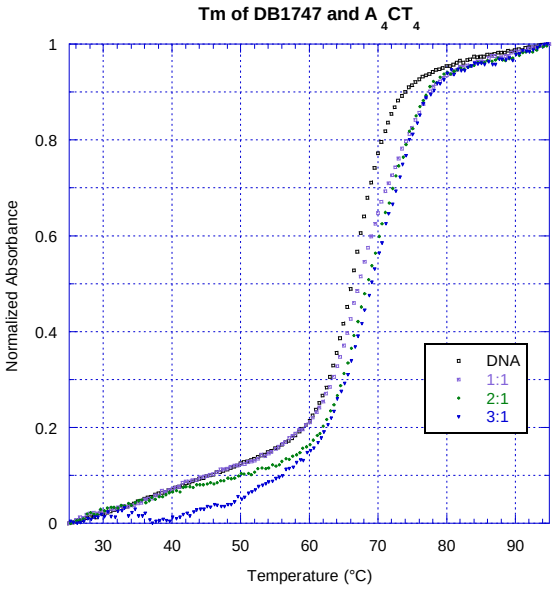

Tm of DB321 and A 4CT4
Normalized Absorbance
Temperature (°C)
DNA
0.5:1
1:1
2:1
3:1

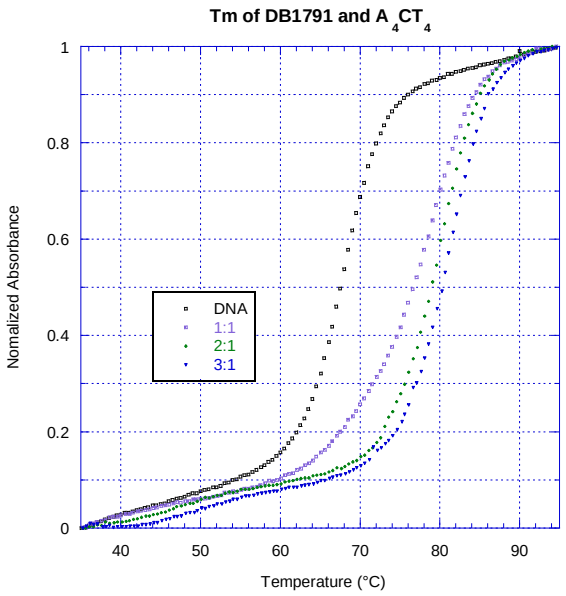

Tm of DB325, DB496, DB2247, DB2260, and A 4CT4
Normalized Absorbance
Temperature (°C)
DNA
DB325 (2:1)
DB496 (2:1)
DB2247 (2:1)
DB2260 (2:1)

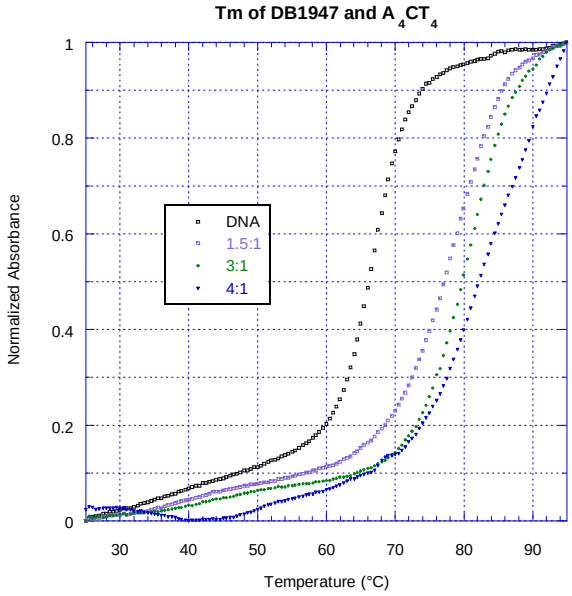

Tm of DB462 and A4CT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

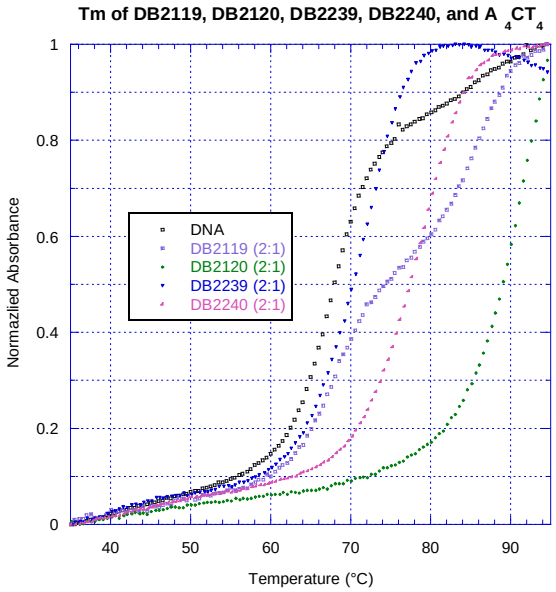

Tm of DB510 and A4CT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

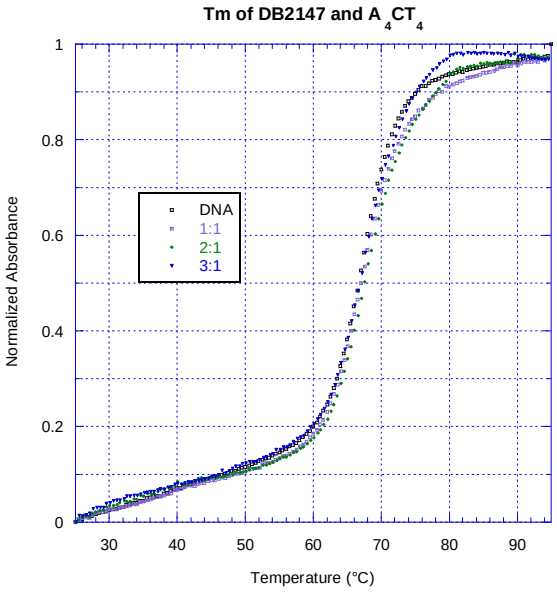

Tm of DB1746 and A4CT4
Normalzed Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

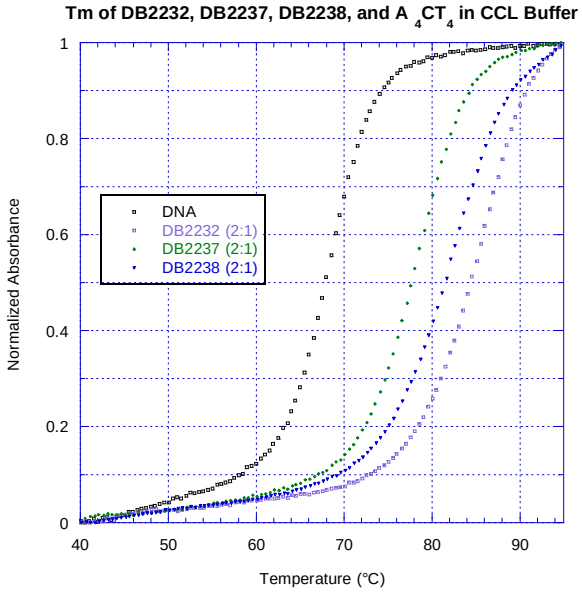

Tm of DB1747 and A4CT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

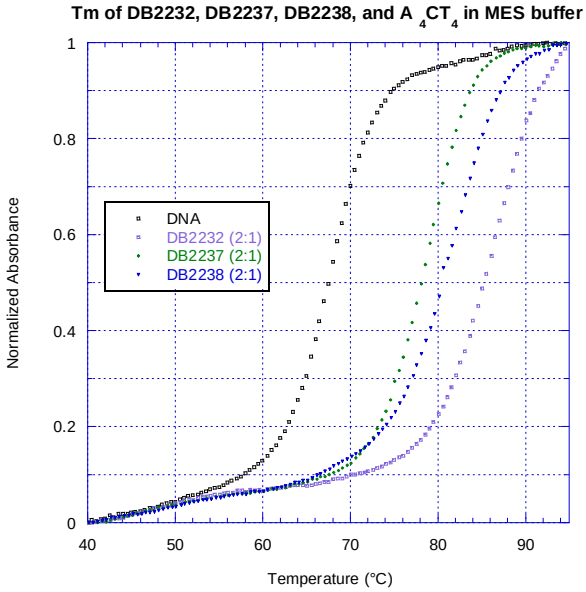

Tm of DB1791 and A4CT4
Nomalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

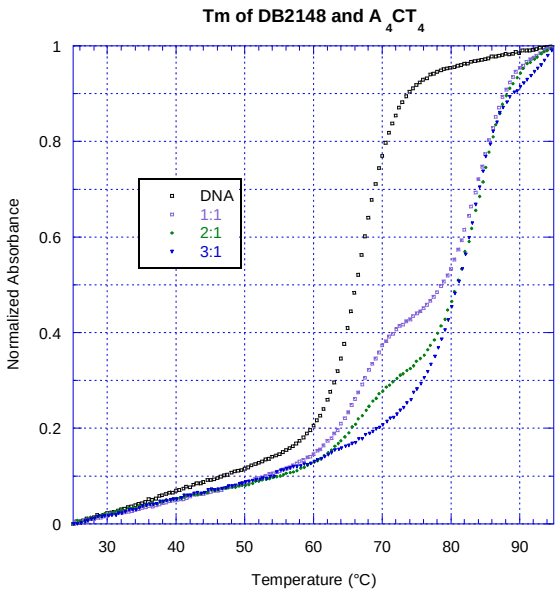

Tm of DB1947 and A4CT4
Normalized Absorbance
Temperature (°C)
DNA
1.5:1
3:1
4:1

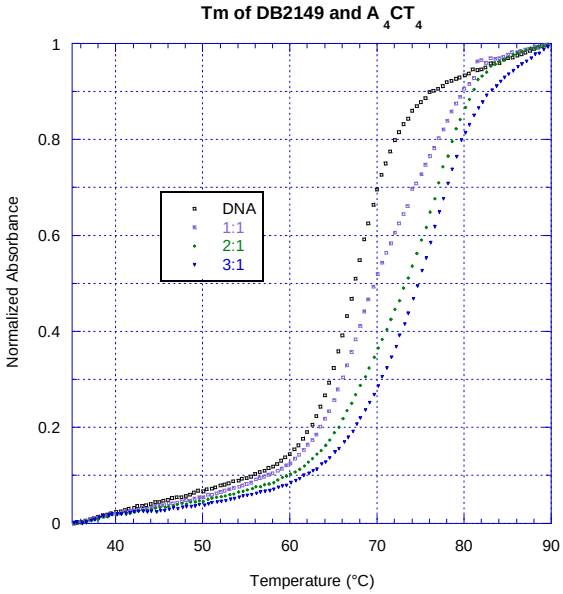

Tm of DB2119, DB2120, DB2239, DB2240, and A4CT4
Normazlied Absorbance
Temperature (°C)
DNA
DB2119 (2:1)
DB2120 (2:1)
DB2239 (2:1)
DB2240 (2:1)

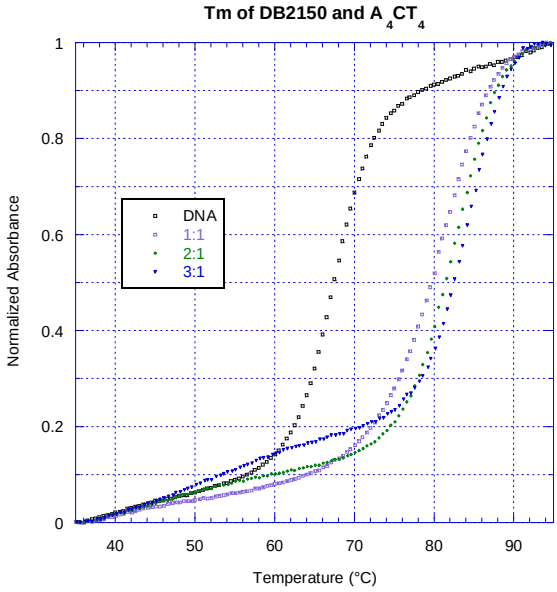

Tm of DB2147 and A4CT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1


Tm of DB2232, DB2237, DB2238, and A4CT4 in CCL Buffer
Normalized Absorbance
Temperature (°C)
DNA
DB2232 (2:1)
DB2237 (2:1)
DB2238 (2:1)

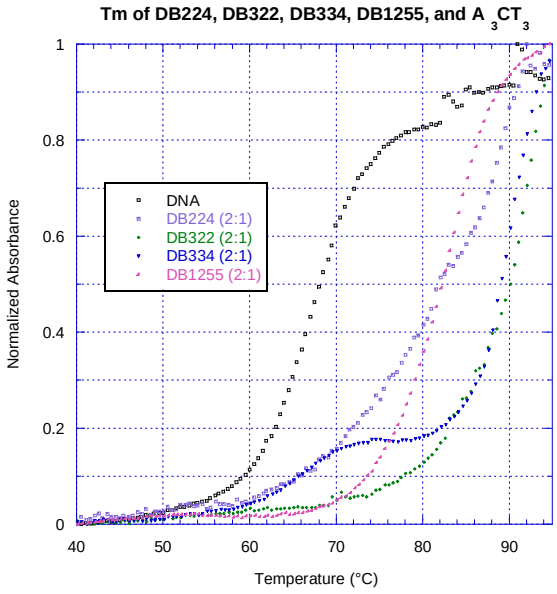

Tm of DB2232, DB2237, DB2238, and A4CT4 in MES buffer
Normalized Absorbance
Temperature (°C)
DNA
DB2232 (2:1)
DB2237 (2:1)
DB2238 (2:1)

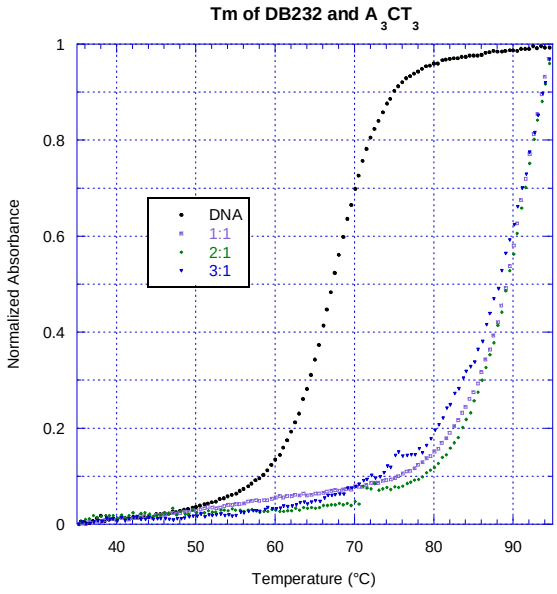

Tm of DB2148 and A4CT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

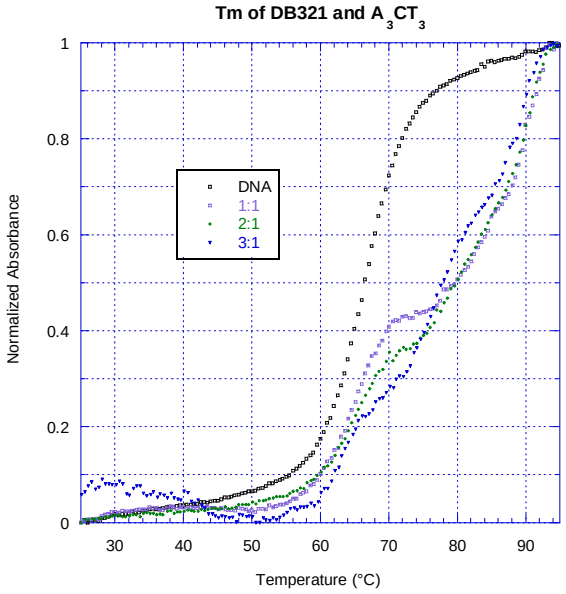

Tm of DB2149 and A4CT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

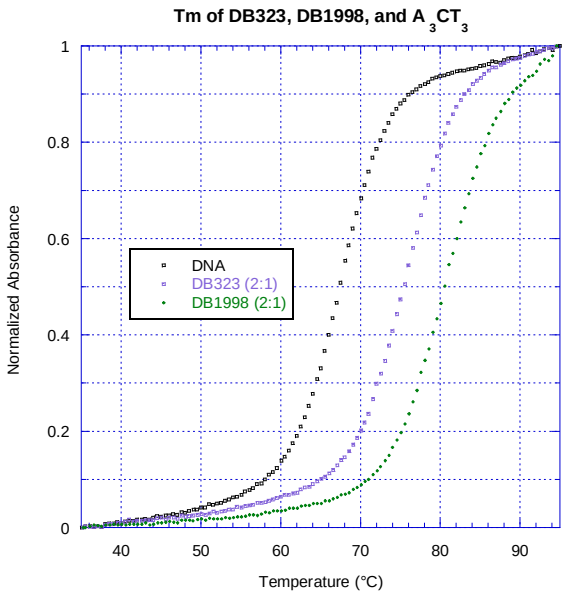

Tm of DB2150 and A4CT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

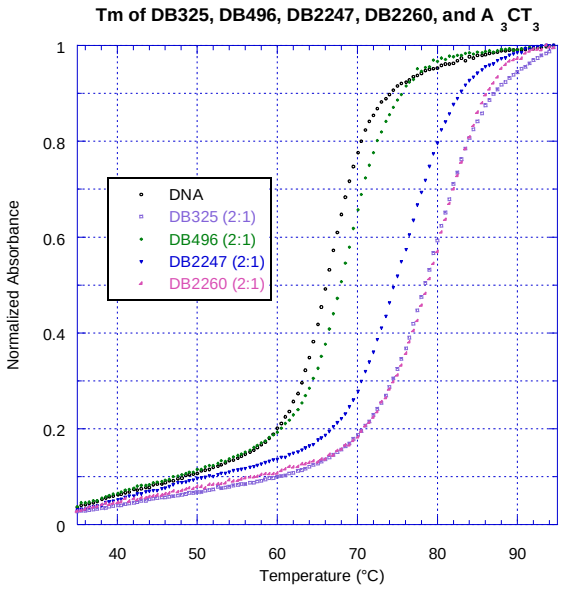
A_3CT_3:

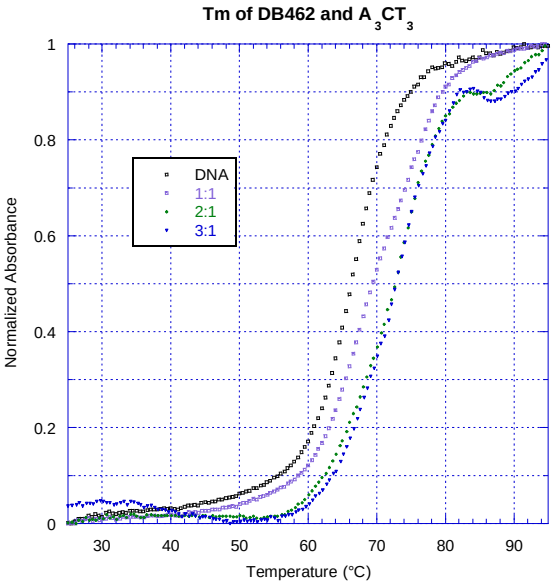

Tm of DB224, DB322, DB334, DB1255, and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB322 (2:1)
DB334 (2:1)
DB1255 (2:1)

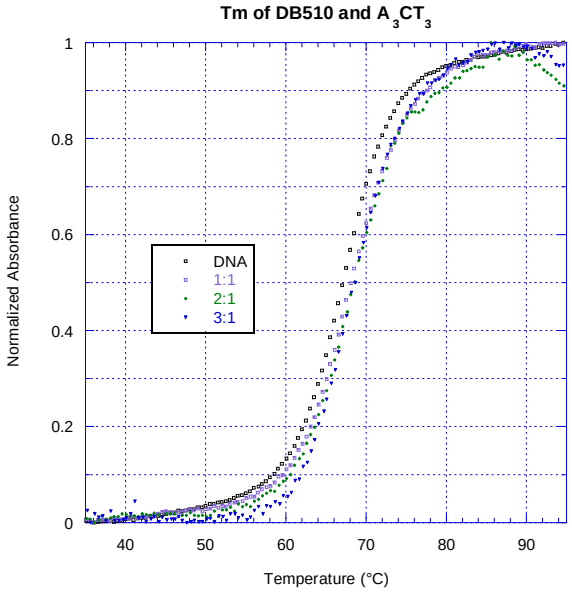

Tm of DB232 and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

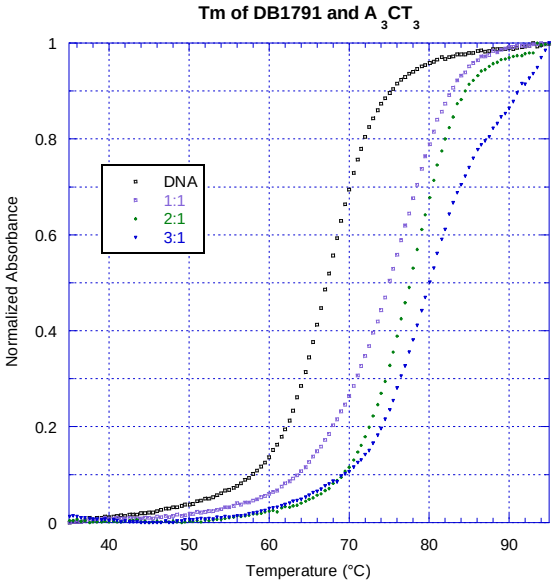

Tm of DB321 and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

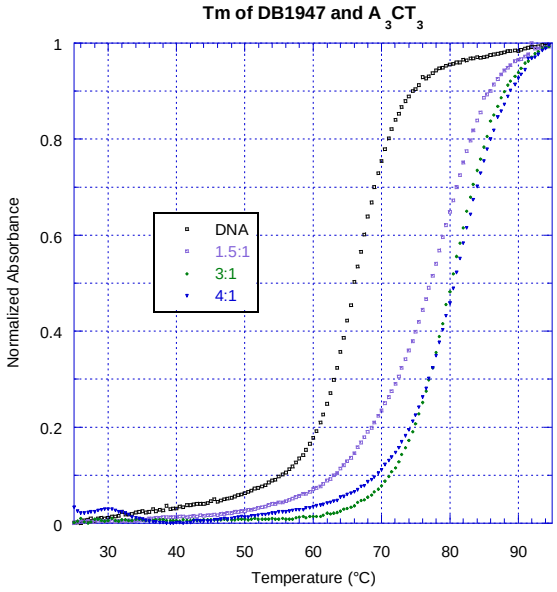

Tm of DB323, DB1998, and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
DB323 (2:1)
DB1998 (2:1)

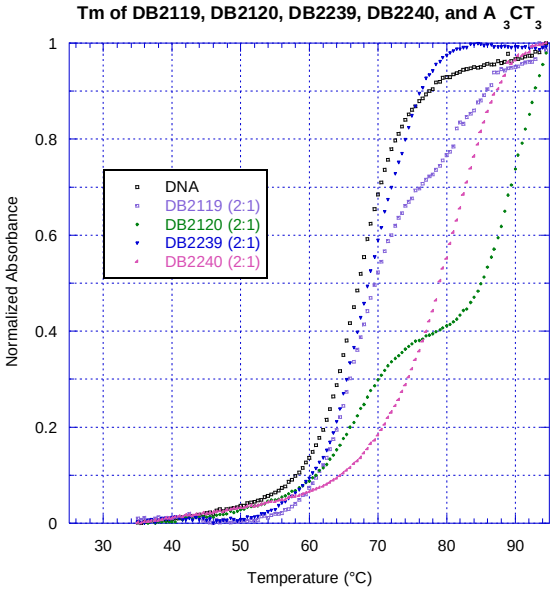

Tm of DB325, DB496, DB2247, DB2260, and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
DB325 (2:1)
DB496 (2:1)
DB2247 (2:1)
DB2260 (2:1)

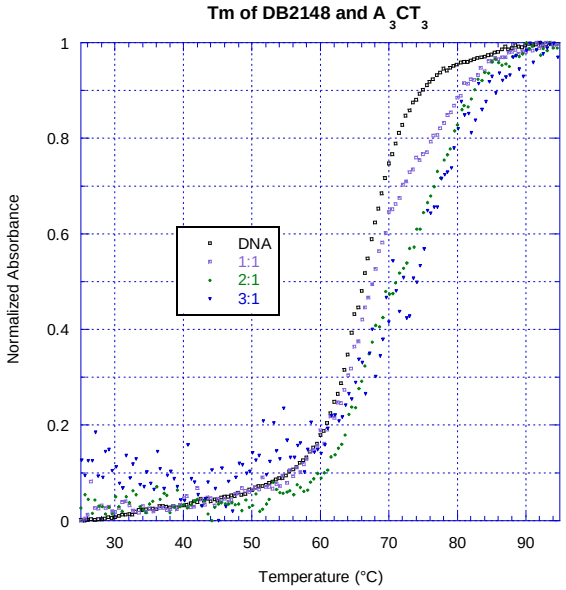

Tm of DB462 and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

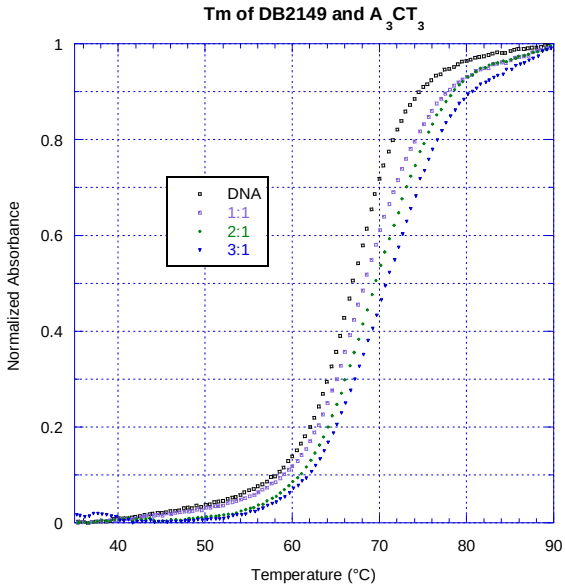

Tm of DB510 and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

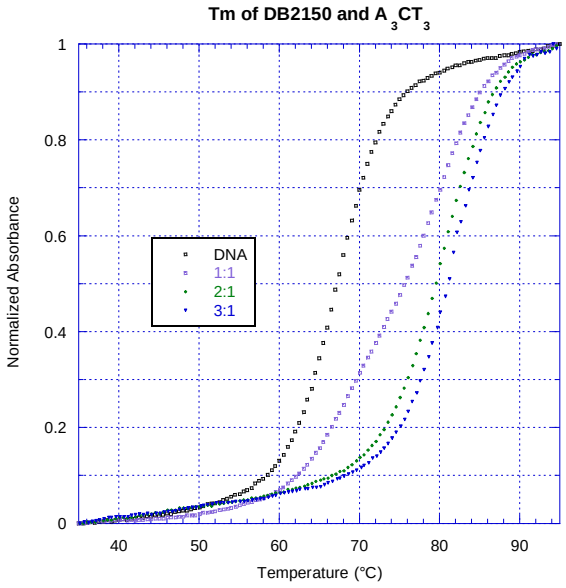

Tm of DB1791 and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

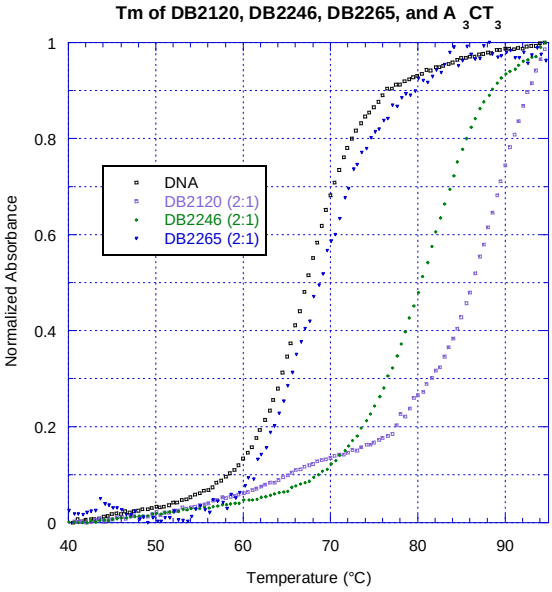

Tm of DB1947 and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
1.5:1
3:1
4:1

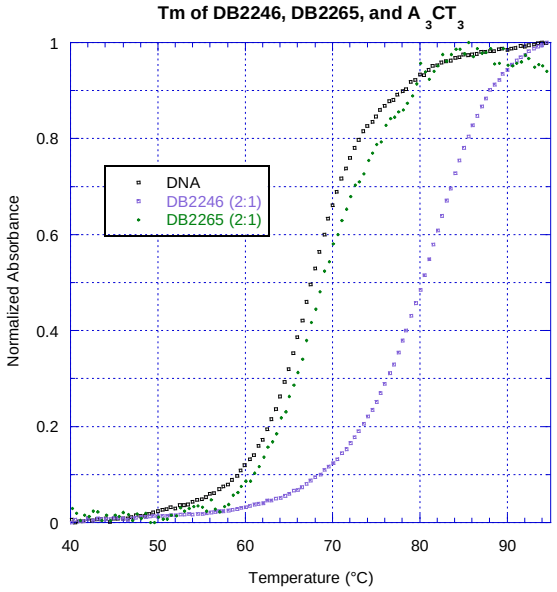

Tm of DB2119, DB2120, DB2239, DB2240, and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
DB2119 (2:1)
DB2120 (2:1)
DB2239 (2:1)
DB2240 (2:1)

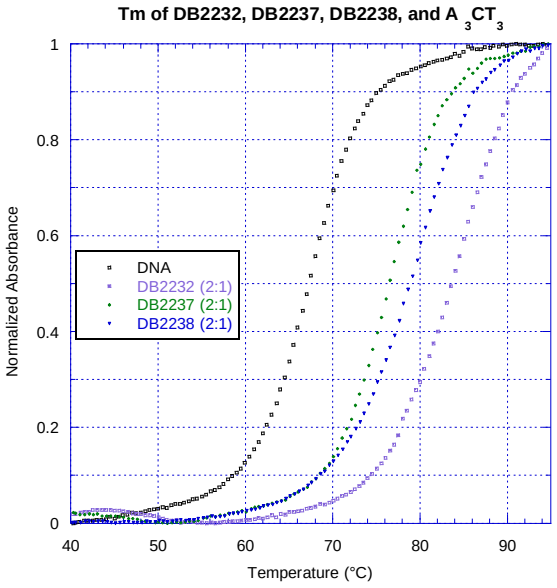

Tm of DB2148 and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1


Tm of DB2149 and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

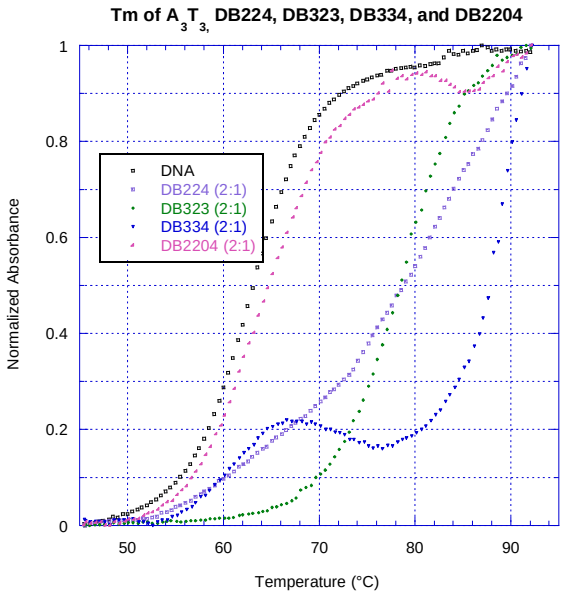

Tm of DB2150 and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

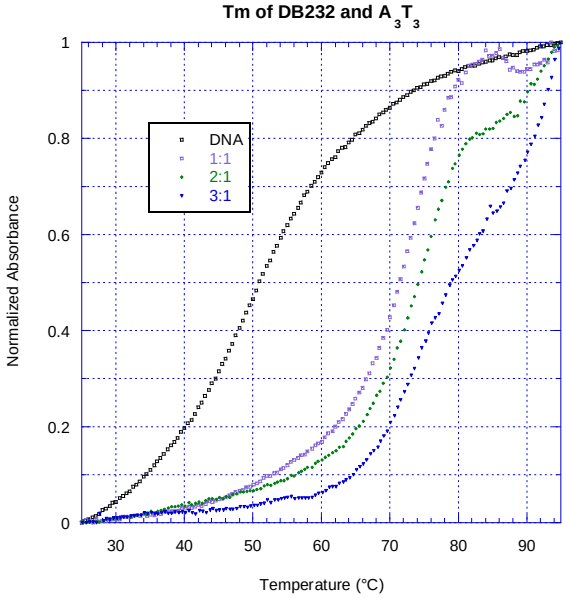

Tm of DB2120, DB2246, DB2265, and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
DB2120 (2:1)
DB2246 (2:1)
DB2265 (2:1)

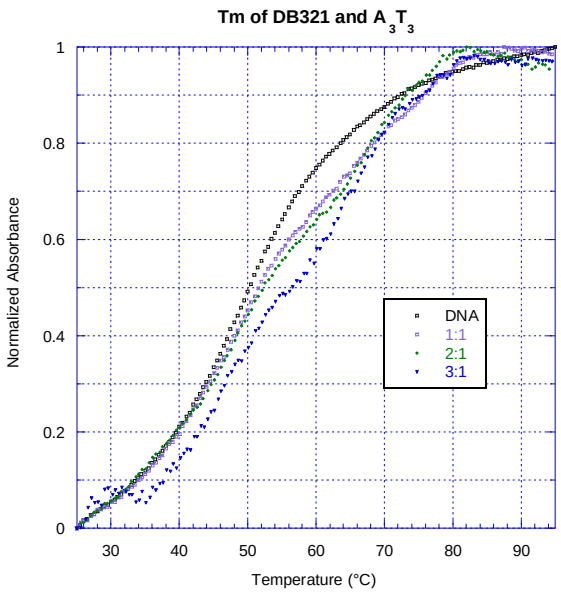

Tm of DB2246, DB2265, and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
DB2246 (2:1)
DB2265 (2:1)

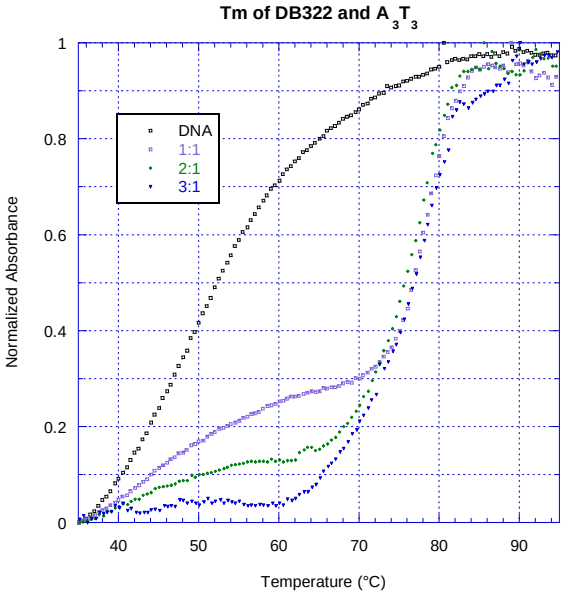

Tm of DB2232, DB2237, DB2238, and A3CT3
Normalized Absorbance
Temperature (°C)
DNA
DB2232 (2:1)
DB2237 (2:1)
DB2238 (2:1)

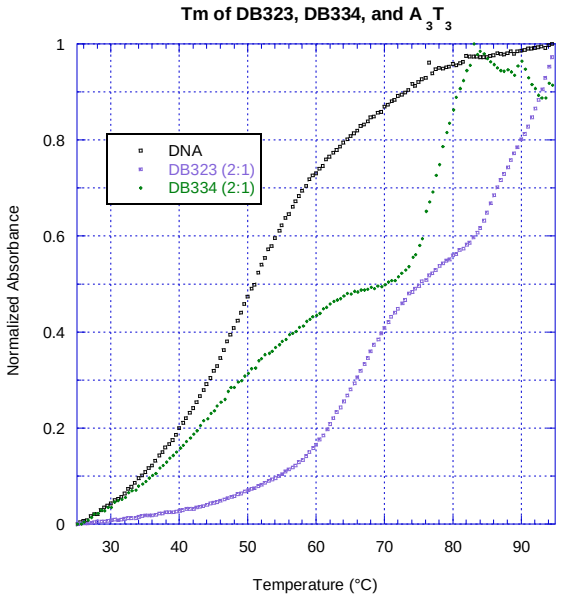
A_3T_3:

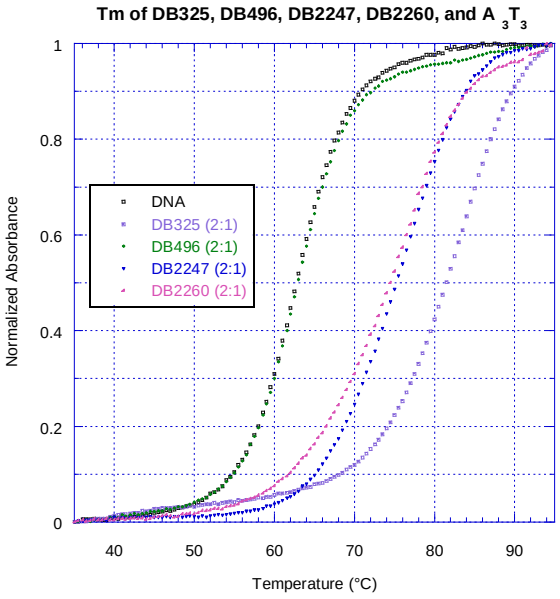

Tm of A3T3, DB224, DB323, DB334, and DB2204
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB323 (2:1)
DB334 (2:1)
DB2204 (2:1)

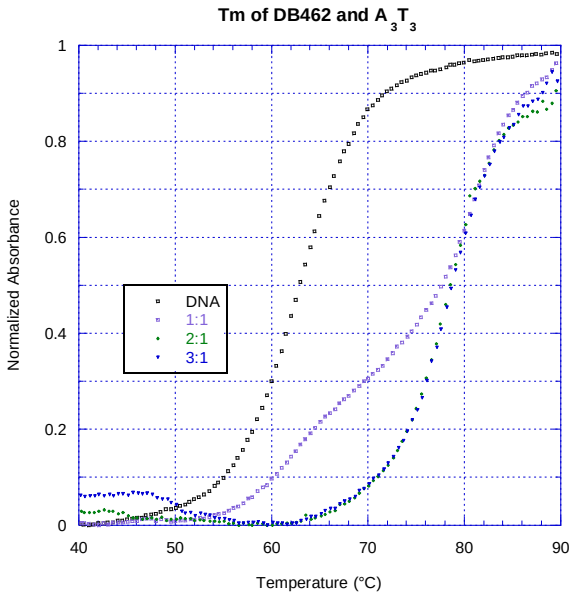

Tm of DB232 and A3T3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

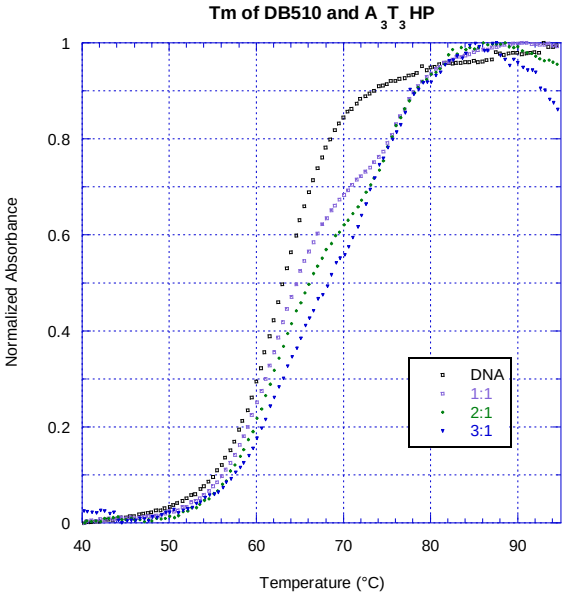

Tm of DB321 and A3T3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

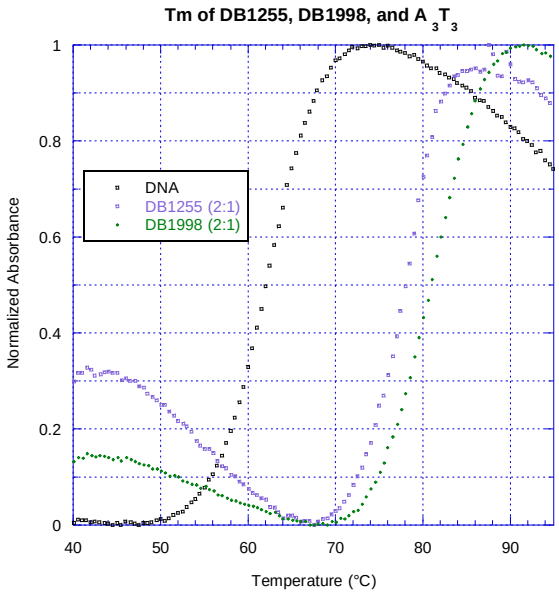

Tm of DB322 and A3T3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

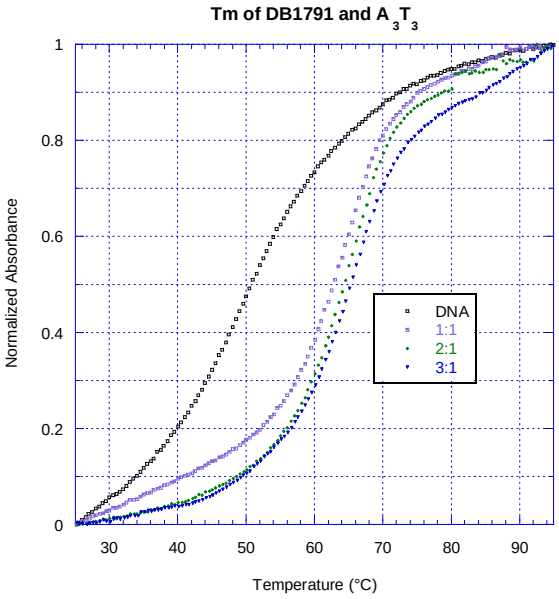

Tm of DB323, DB334, and A3T3
Normalized Absorbance
Temperature (°C)
DNA
DB323 (2:1)
DB334 (2:1)

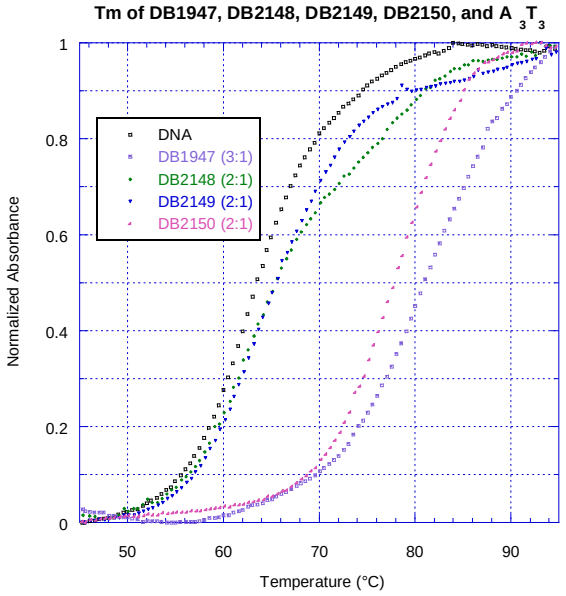

Tm of DB325, DB496, DB2247, DB2260, and A3T3
Normalized Absorbance
Temperature (°C)
DNA
DB325 (2:1)
DB496 (2:1)
DB2247 (2:1)
DB2260 (2:1)

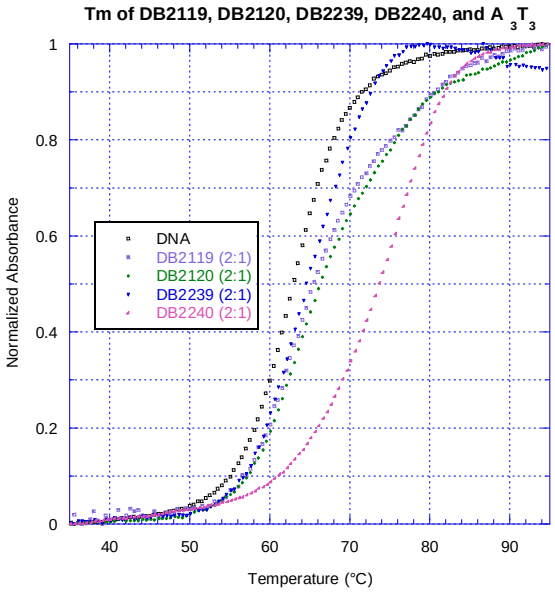

Tm of DB462 and A3T3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

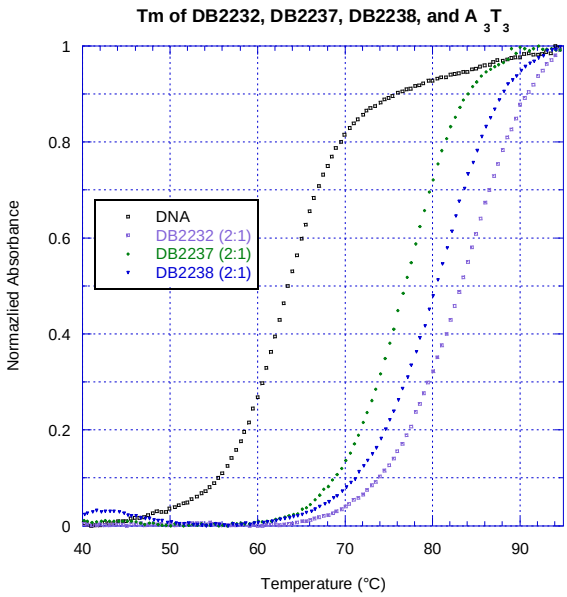

Tm of DB510 and A3T3 HP
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1


Tm of DB1255, DB1998, and A3T3
Normalized Absorbance
Temperature (°C)
DNA
DB1255 (2:1)
DB1998 (2:1)

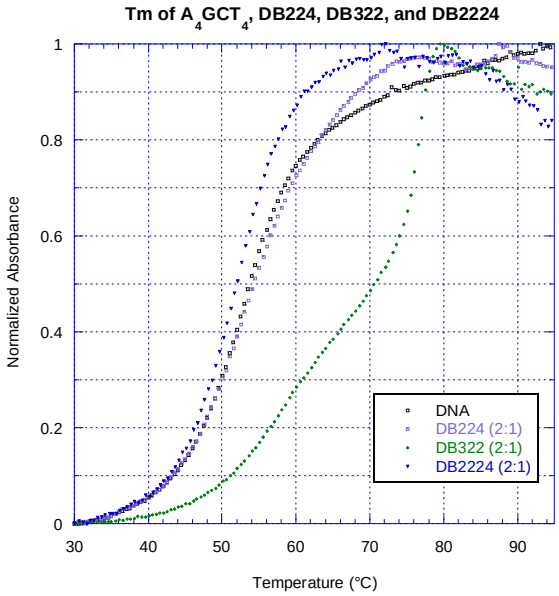

Tm of DB1791 and A3T3
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

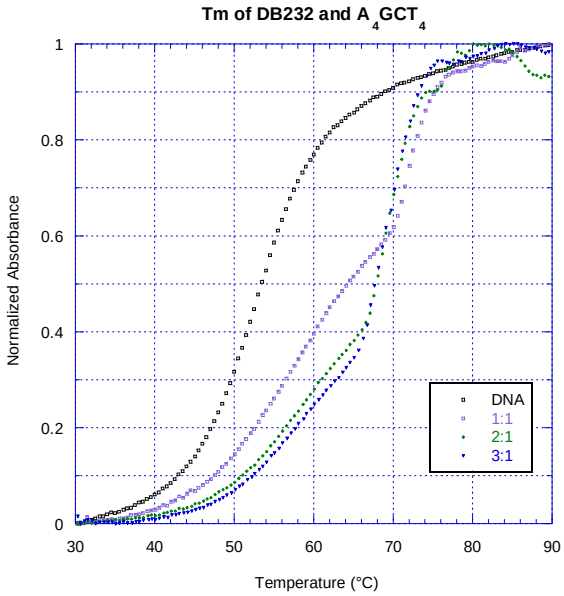

Tm of DB1947, DB2148, DB2149, DB2150, and A3T3
Normalized Absorbance
Temperature (°C)
DNA
DB1947 (3:1)
DB2148 (2:1)
DB2149 (2:1)
DB2150 (2:1)

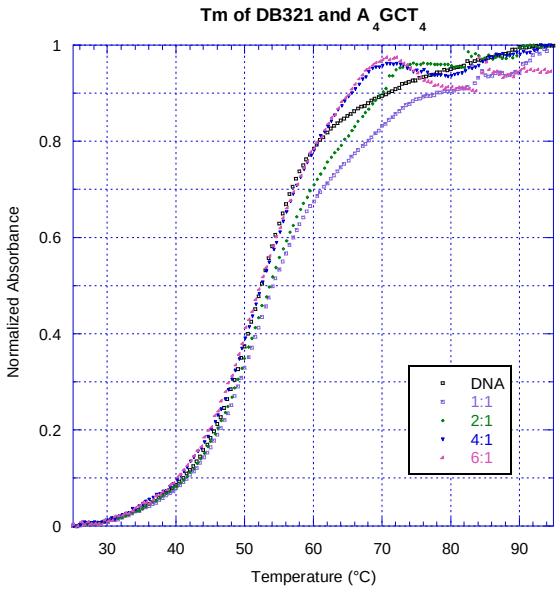

Tm of DB2119, DB2120, DB2239, DB2240, and A3T3
Normalized Absorbance
Temperature (°C)
DNA
DB2119 (2:1)
DB2120 (2:1)
DB2239 (2:1)
DB2240 (2:1)

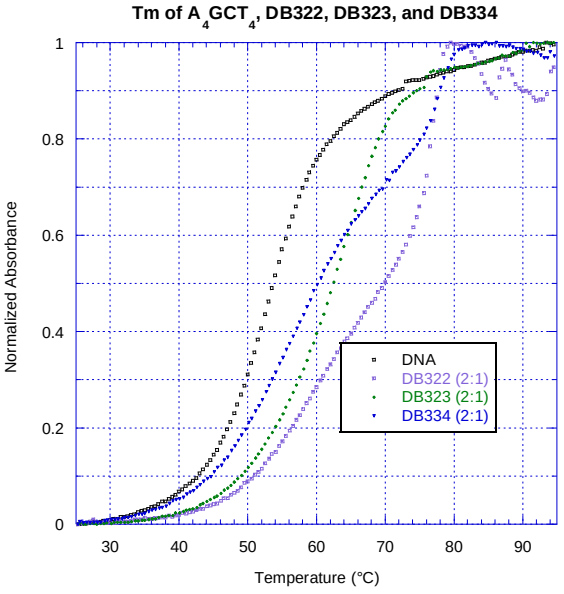

Tm of DB2232, DB2237, DB2238, and A3T3
Normazlied Absorbance
Temperature (°C)
DNA
DB2232 (2:1)
DB2237 (2:1)
DB2238 (2:1)

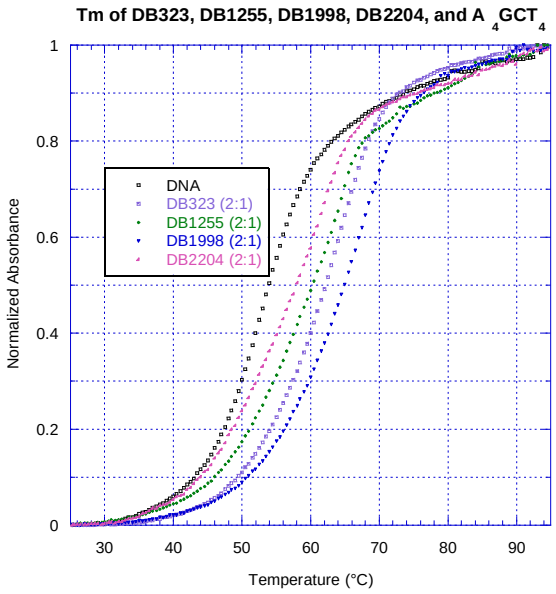
A_4GCT_4:

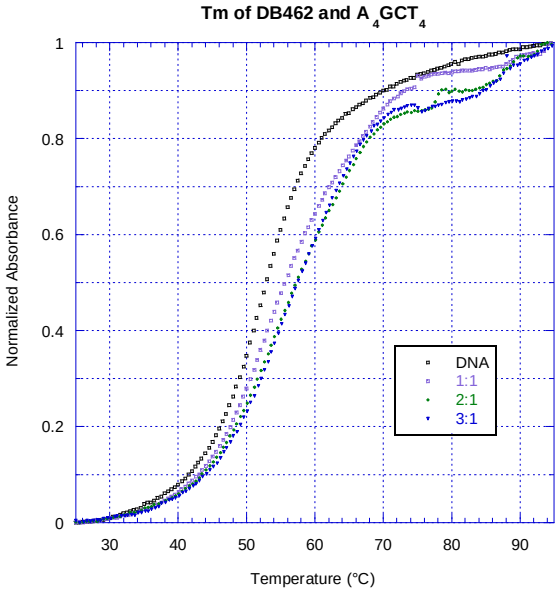

Tm of A4GCT4, DB224, DB322, and DB2224
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB322 (2:1)
DB2224 (2:1)

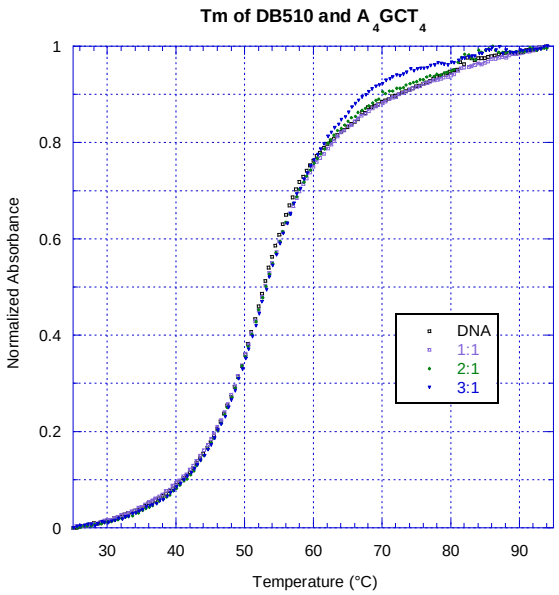

Tm of DB232 and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

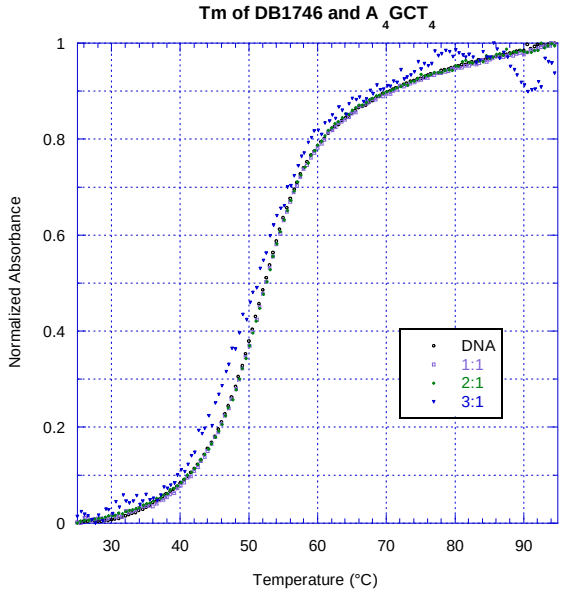

Tm of DB321 and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
4:1
6:1

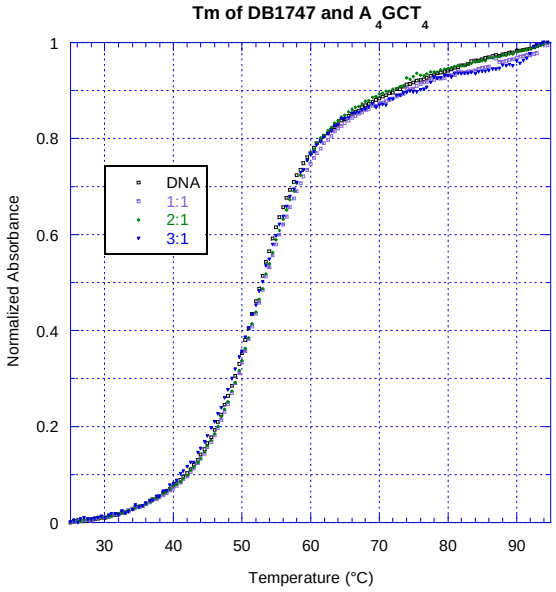

Tm of A4GCT4, DB322, DB323, and DB334
Normalized Absorbance
Temperature (°C)
DNA
DB322 (2:1)
DB323 (2:1)
DB334 (2:1)

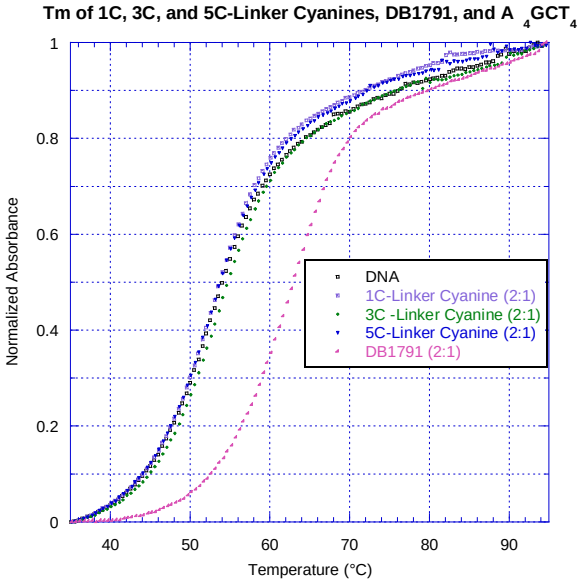

Tm of DB323, DB1255, DB1998, DB2204, and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
DB323 (2:1)
DB1255 (2:1)
DB1998 (2:1)
DB2204 (2:1)

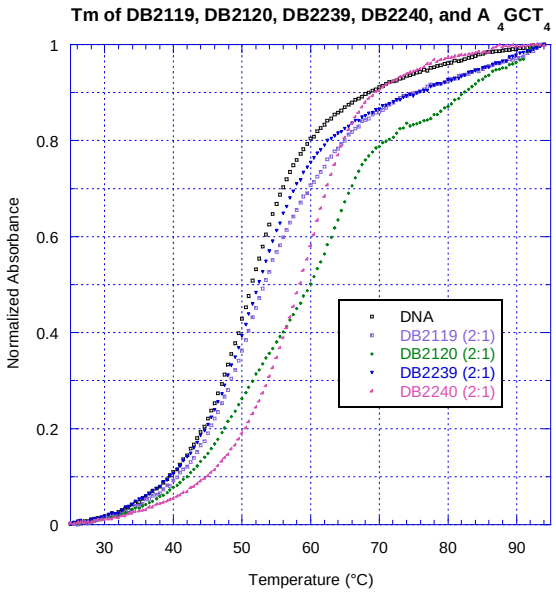

Tm of DB462 and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

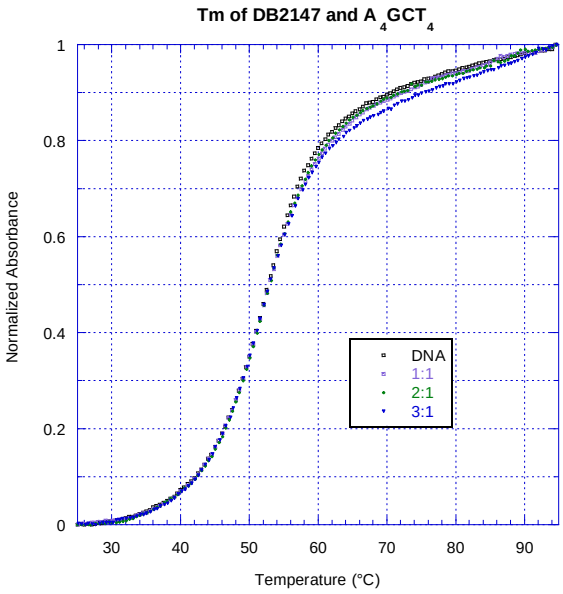

Tm of DB510 and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

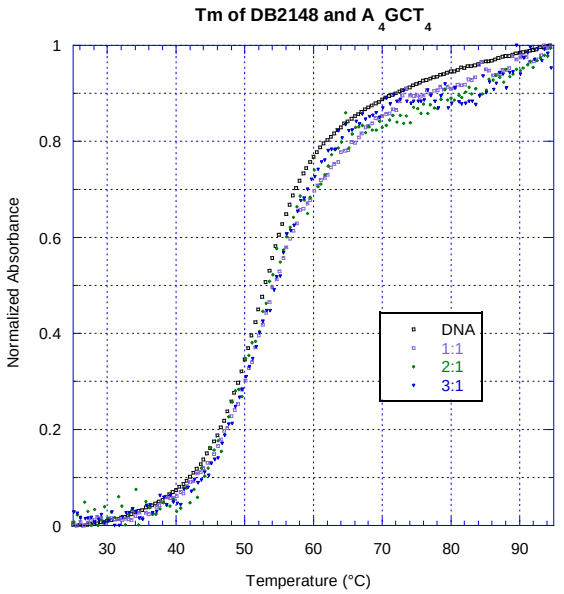

Tm of DB1746 and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

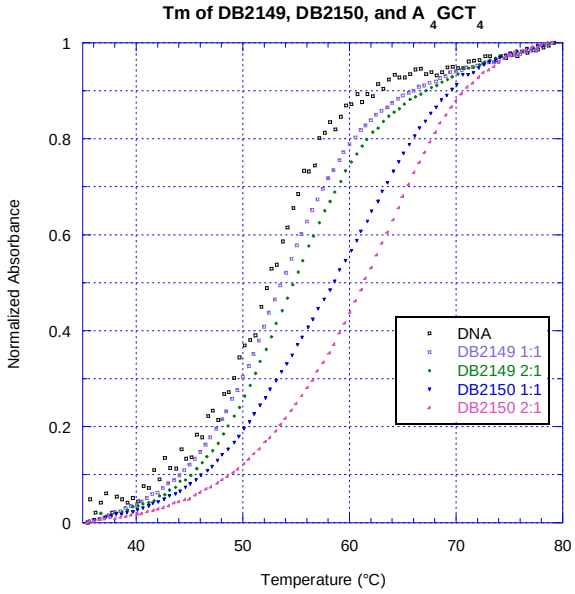

Tm of DB1747 and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

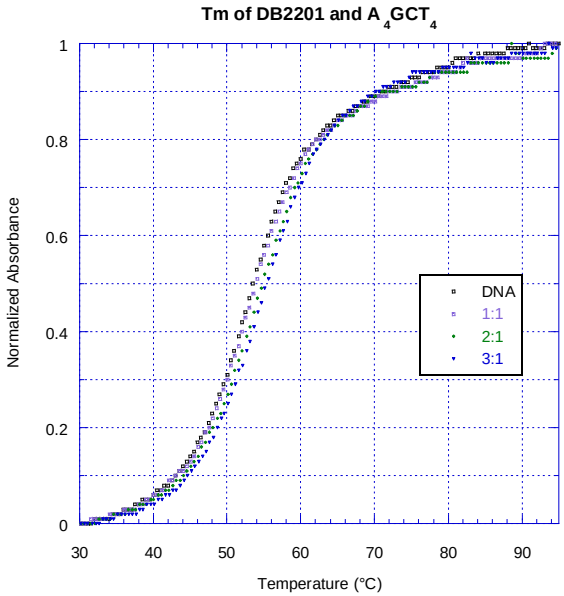

Tm of 1C, 3C, and 5C-Linker Cyanines, DB1791, and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
1C-Linker Cyanine (2:1)
3C -Linker Cyanine (2:1)
5C-Linker Cyanine (2:1)
DB1791 (2:1)

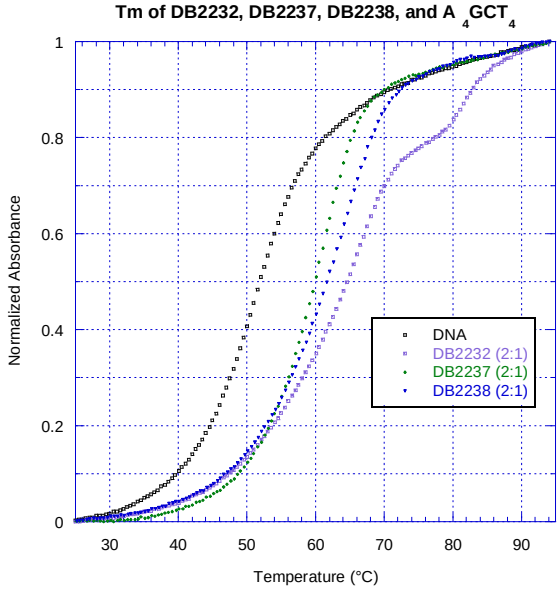

Tm of DB2119, DB2120, DB2239, DB2240, and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
DB2119 (2:1)
DB2120 (2:1)
DB2239 (2:1)
DB2240 (2:1)


Tm of DB2147 and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

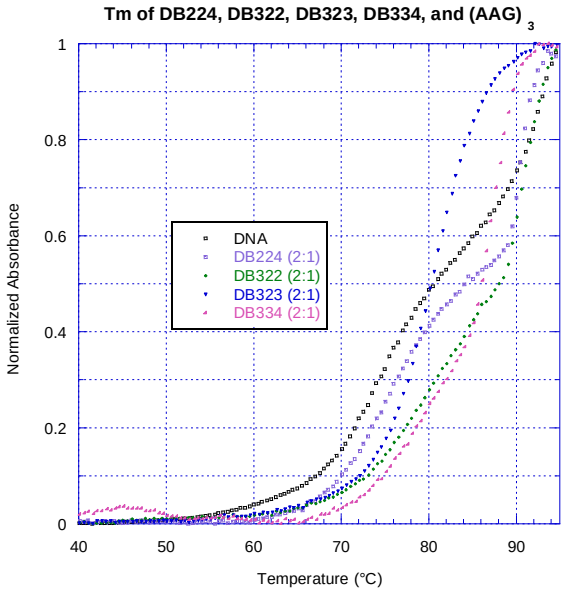

Tm of DB2148 and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

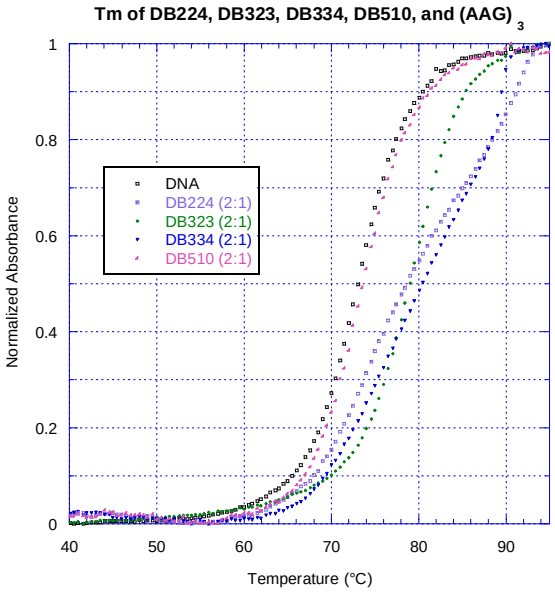

Tm of DB2149, DB2150, and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
DB2149 1:1
DB2149 2:1
DB2150 1:1
DB2150 2:1

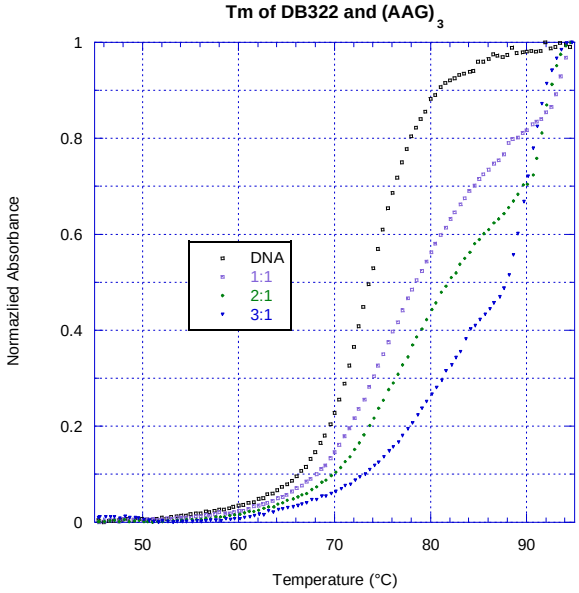

Tm of DB2201 and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

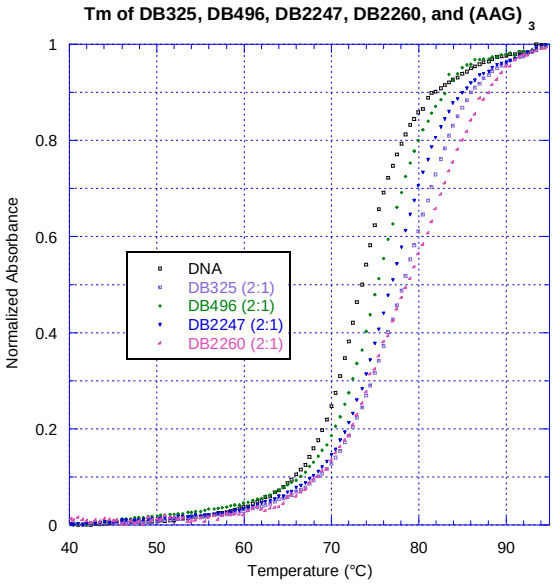

Tm of DB2232, DB2237, DB2238, and A4GCT4
Normalized Absorbance
Temperature (°C)
DNA
DB2232 (2:1)
DB2237 (2:1)
DB2238 (2:1)

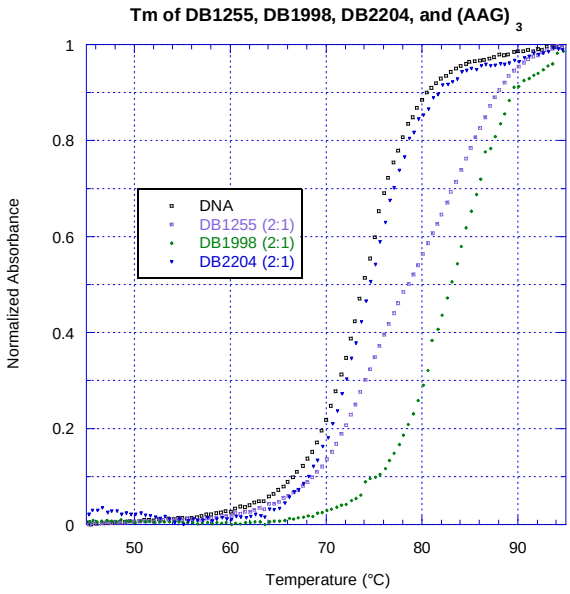
$(AAG)_3$:

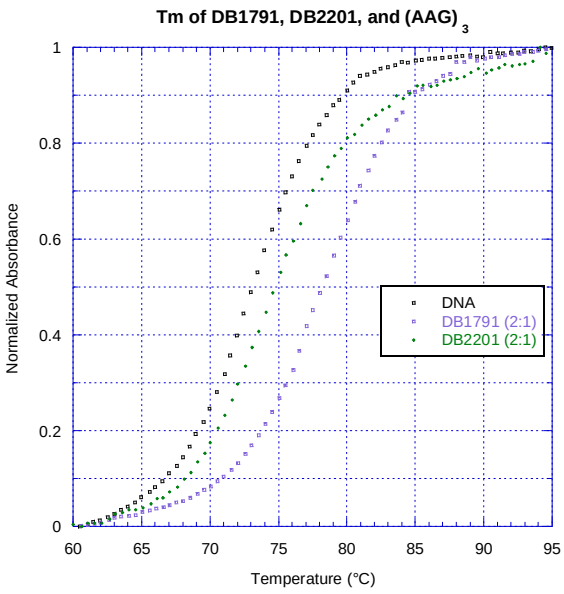

Tm of DB224, DB322, DB323, DB334, and (AAG) 3
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB322 (2:1)
DB323 (2:1)
DB334 (2:1)

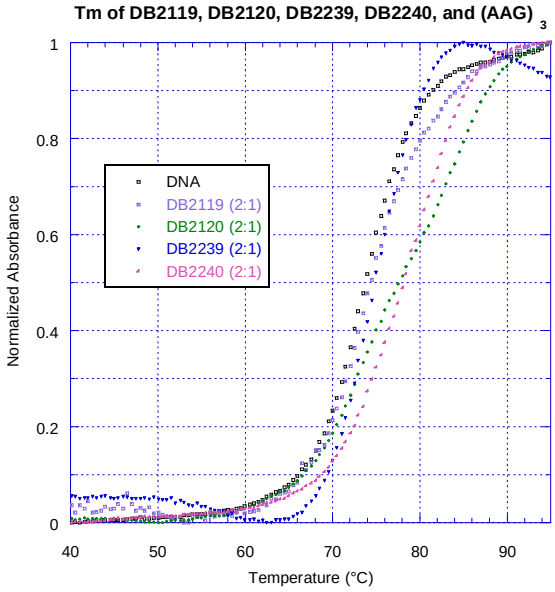

Tm of DB224, DB323, DB334, DB510, and (AAG) 3
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB323 (2:1)
DB334 (2:1)
DB510 (2:1)

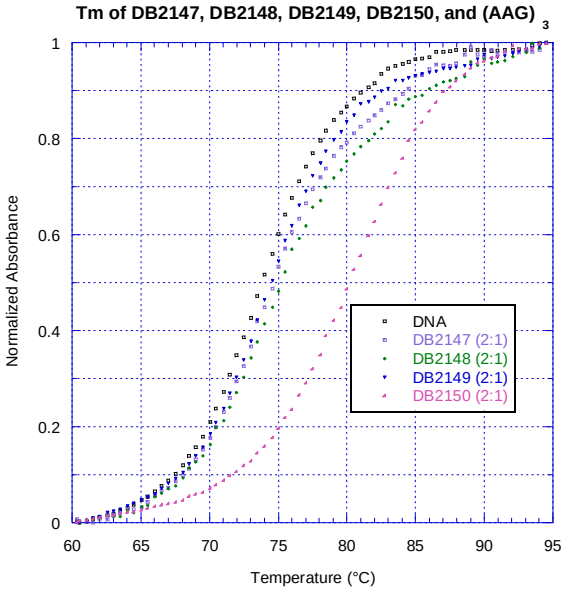

Tm of DB322 and (AAG) 3
Normazlied Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

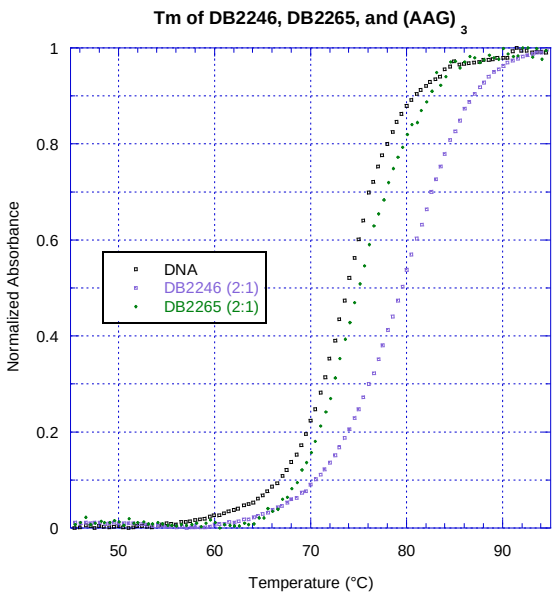

Tm of DB325, DB496, DB2247, DB2260, and (AAG) 3
Normalized Absorbance
Temperature (°C)
DNA
DB325 (2:1)
DB496 (2:1)
DB2247 (2:1)
DB2260 (2:1)

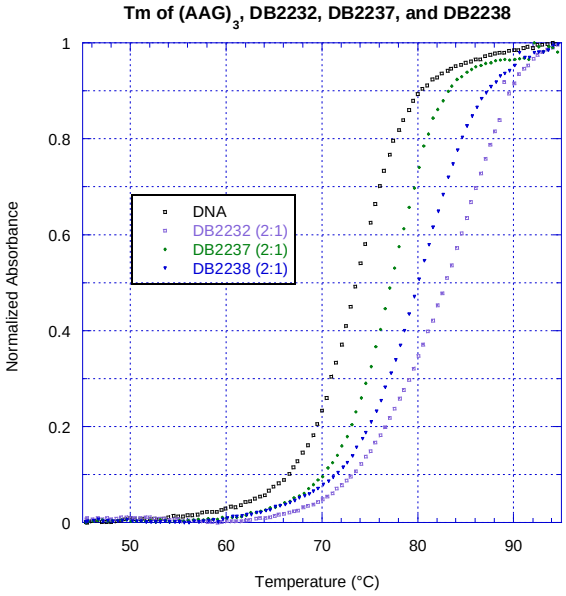

Tm of DB1255, DB1998, DB2204, and (AAG) 3
Normalized Absorbance
Temperature (°C)
DNA
DB1255 (2:1)
DB1998 (2:1)
DB2204 (2:1)


Tm of DB1791, DB2201, and (AAG) 3
Normalized Absorbance
Temperature (°C)
DNA
DB1791 (2:1)
DB2201 (2:1)

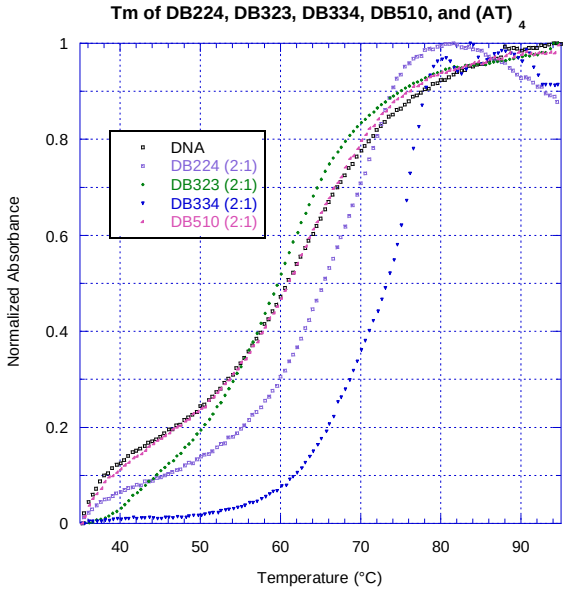

Tm of DB2119, DB2120, DB2239, DB2240, and (AAG) 3
Normalized Absorbance
Temperature (°C)
DNA
DB2119 (2:1)
DB2120 (2:1)
DB2239 (2:1)
DB2240 (2:1)

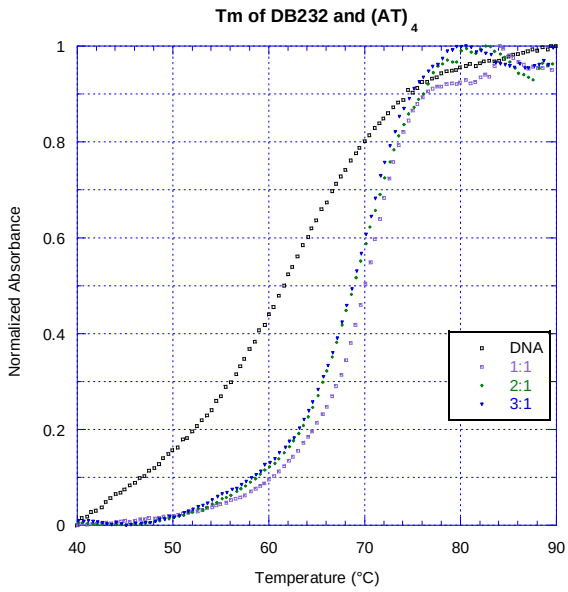

Tm of DB2147, DB2148, DB2149, DB2150, and (AAG) 3
Normalized Absorbance
Temperature (°C)
DNA
DB2147 (2:1)
DB2148 (2:1)
DB2149 (2:1)
DB2150 (2:1)

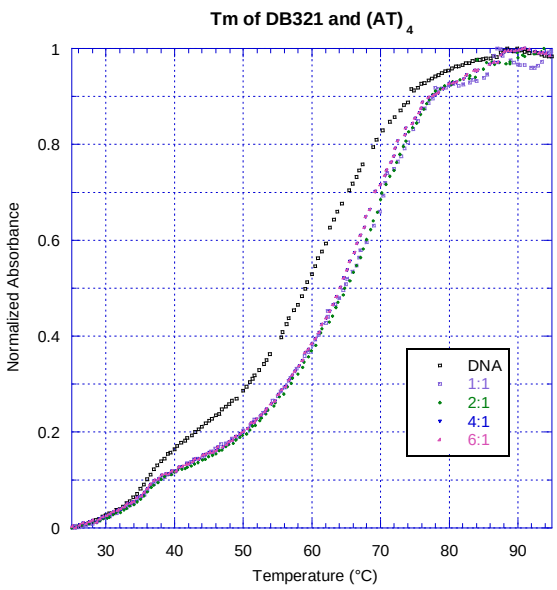

Tm of DB2246, DB2265, and (AAG) 3
Normalized Absorbance
Temperature (°C)
DNA
DB2246 (2:1)
DB2265 (2:1)

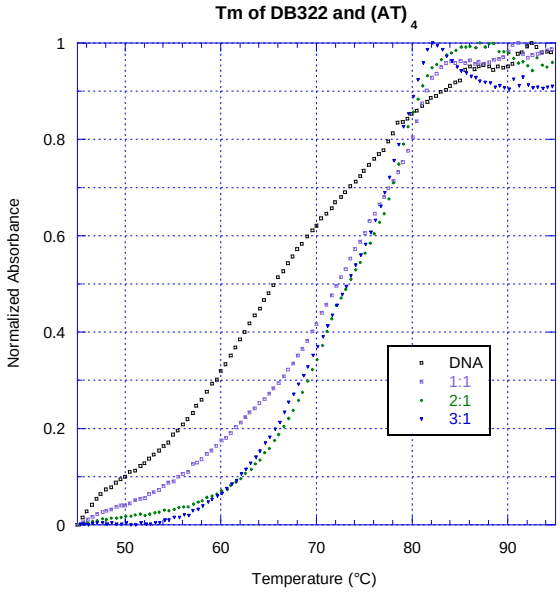

Tm of (AAG)3, DB2232, DB2237, and DB2238
Normalized Absorbance
Temperature (°C)
DNA
DB2232 (2:1)
DB2237 (2:1)
DB2238 (2:1)

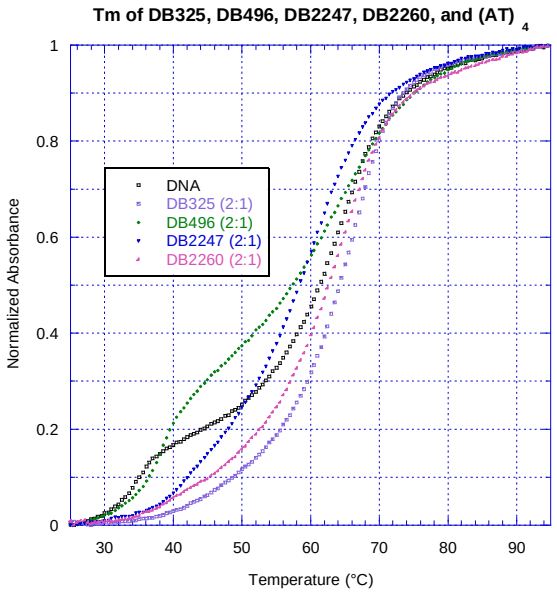
$(AT)_4$:

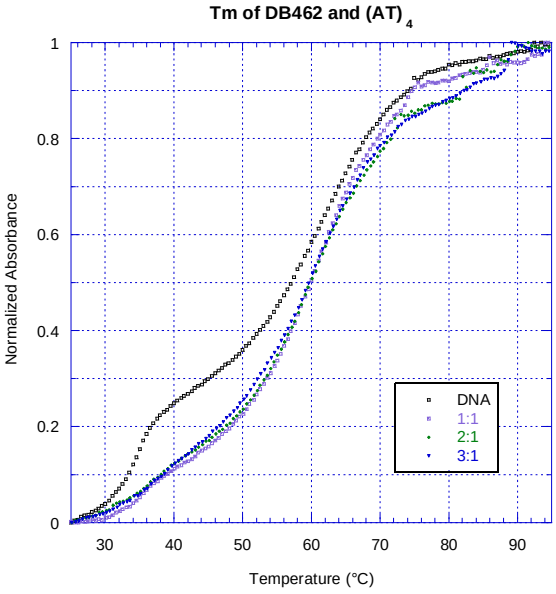

Tm of DB224, DB323, DB334, DB510, and (AT) 4
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB323 (2:1)
DB334 (2:1)
DB510 (2:1)

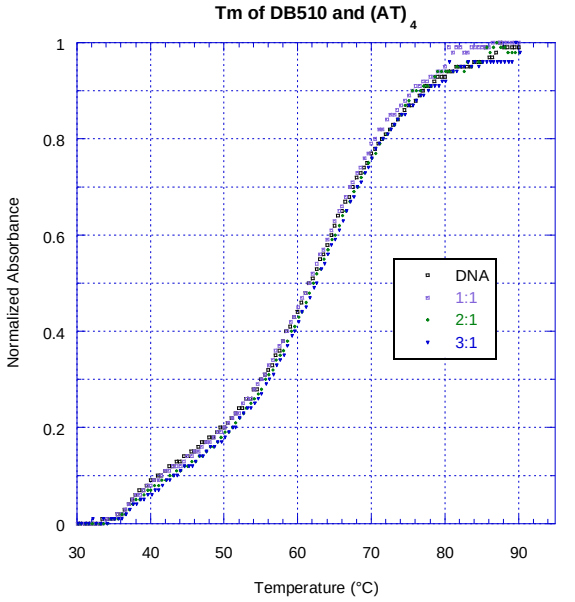

Tm of DB232 and (AT) 4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

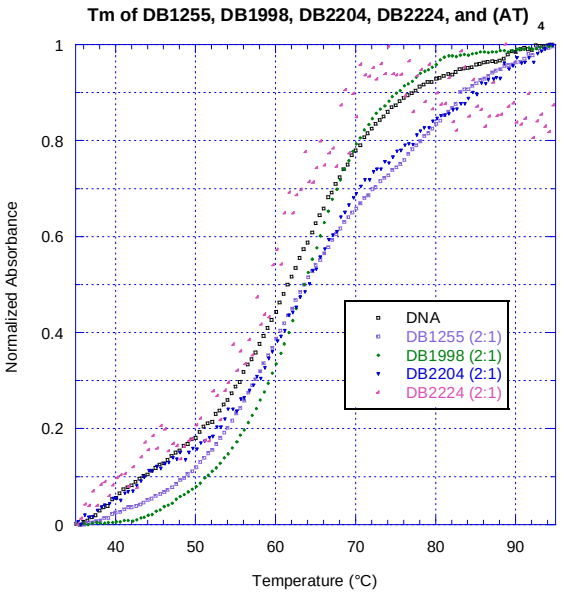

Tm of DB321 and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
4:1
6:1

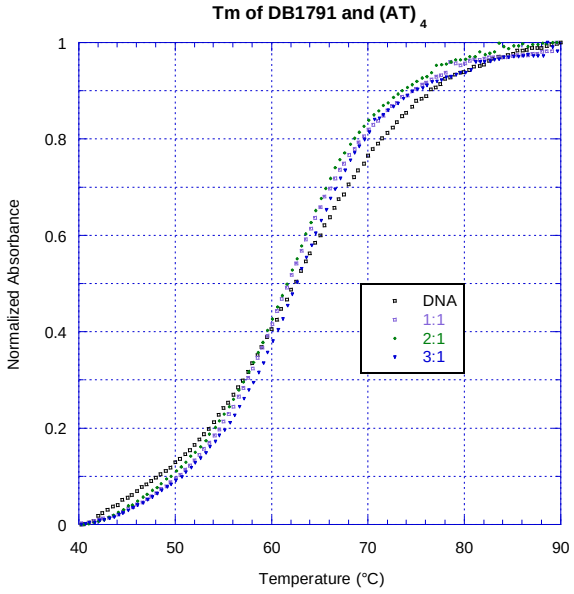

Tm of DB322 and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

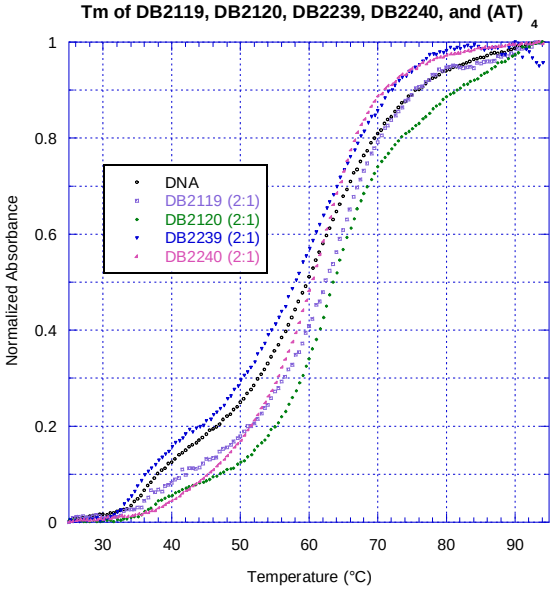

Tm of DB325, DB496, DB2247, DB2260, and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
DB325 (2:1)
DB496 (2:1)
DB2247 (2:1)
DB2260 (2:1)

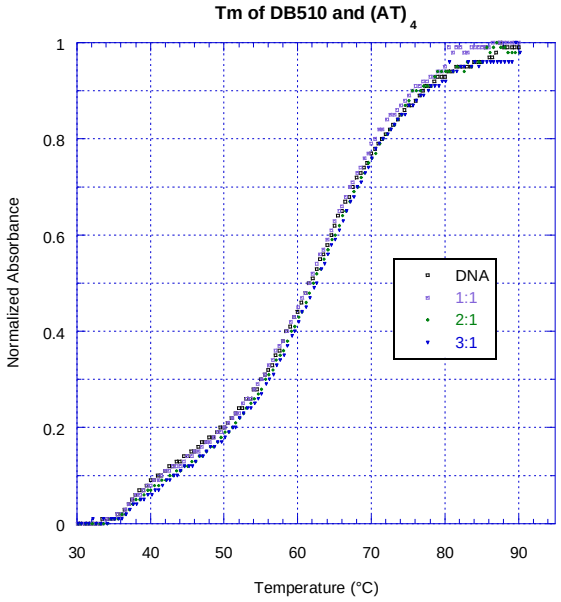

Tm of DB462 and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

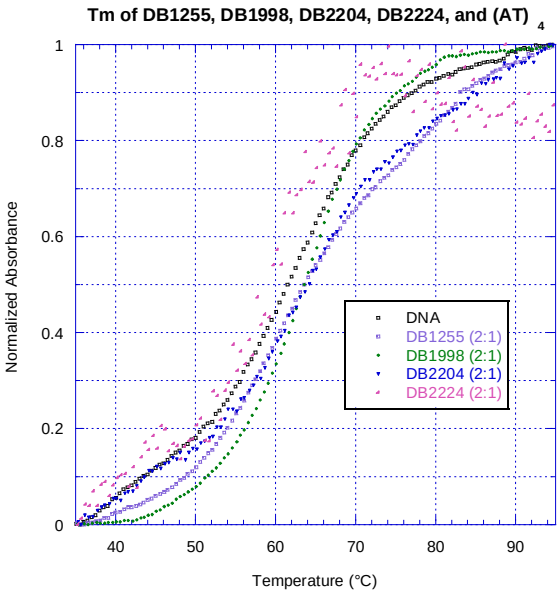

Tm of DB510 and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

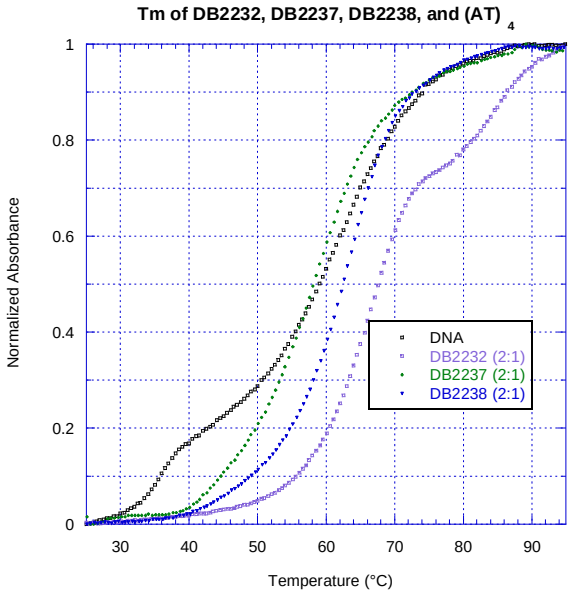

Tm of DB1255, DB1998, DB2204, DB2224, and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
DB1255 (2:1)
DB1998 (2:1)
DB2204 (2:1)
DB2224 (2:1)

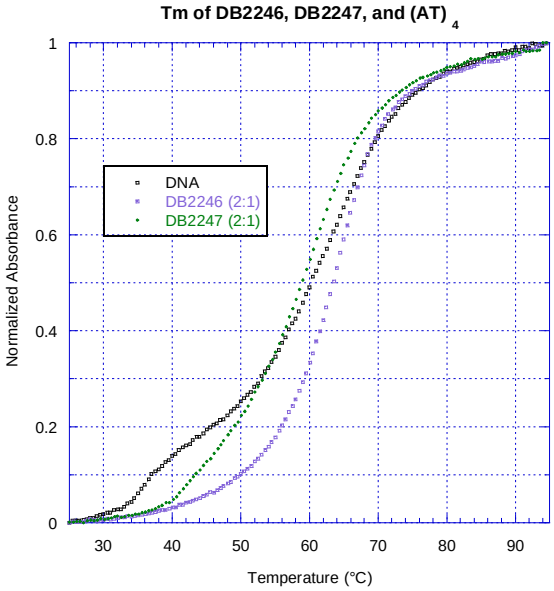

Tm of DB1791 and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1


Tm of DB2119, DB2120, DB2239, DB2240, and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
DB2119 (2:1)
DB2120 (2:1)
DB2239 (2:1)
DB2240 (2:1)

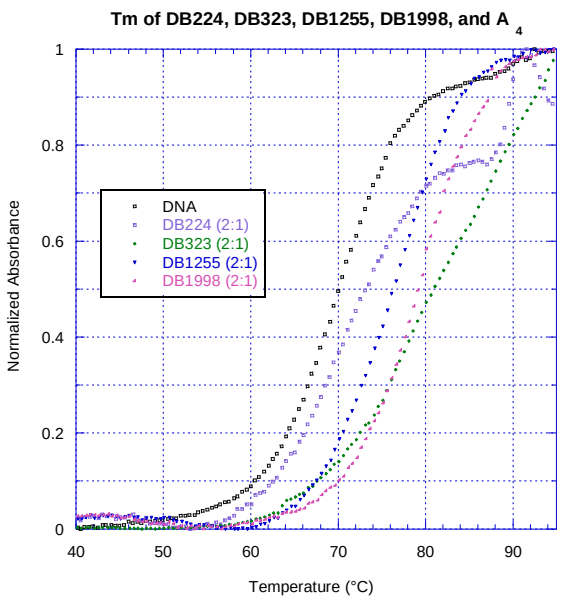

Tm of DB510 and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

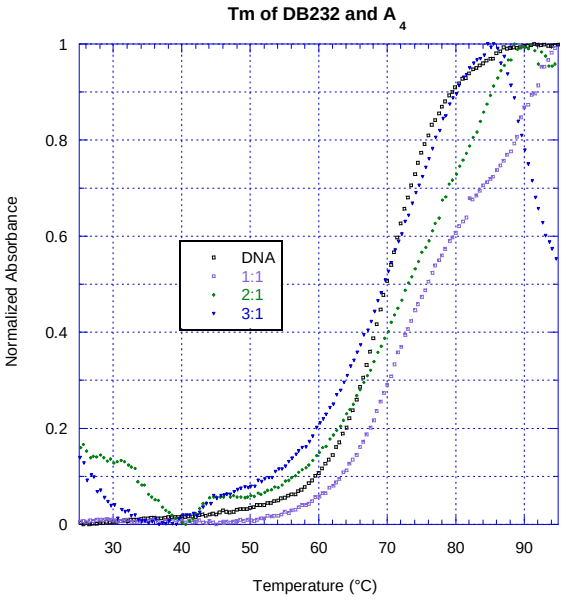

Tm of DB1255, DB1998, DB2204, DB2224, and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
DB1255 (2:1)
DB1998 (2:1)
DB2204 (2:1)
DB2224 (2:1)

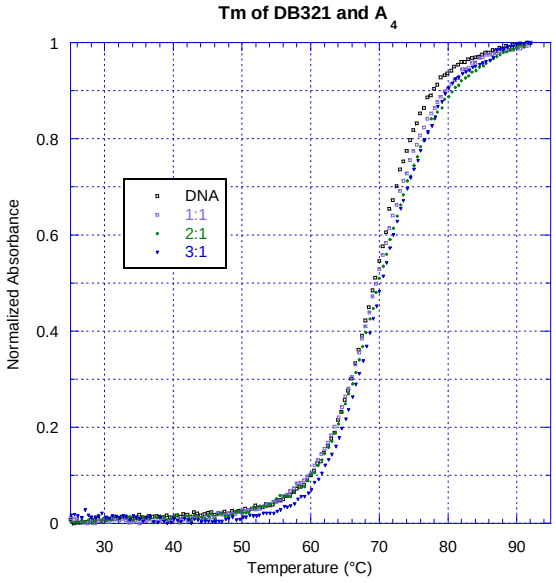

Tm of DB2232, DB2237, DB2238, and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
DB2232 (2:1)
DB2237 (2:1)
DB2238 (2:1)

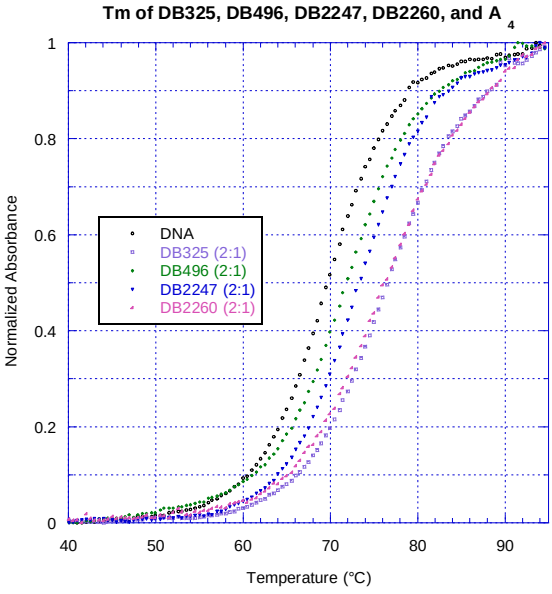

Tm of DB2246, DB2247, and (AT)4
Normalized Absorbance
Temperature (°C)
DNA
DB2246 (2:1)
DB2247 (2:1)

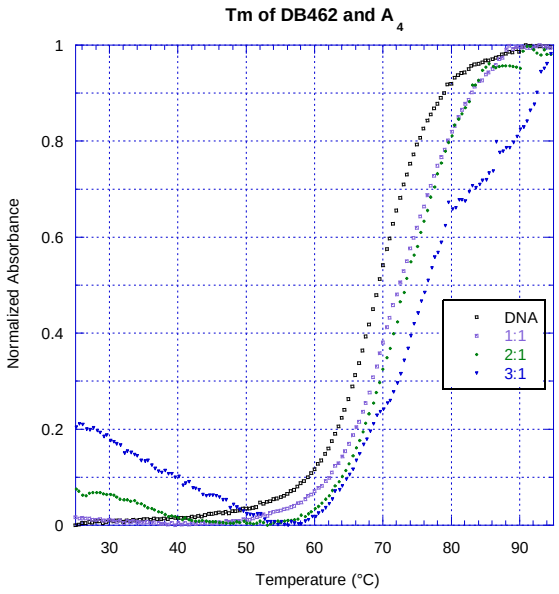
A_4:

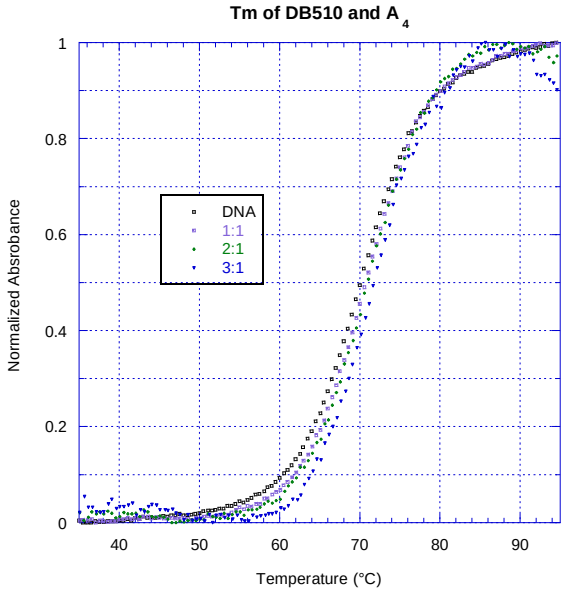

Tm of DB224, DB323, DB1255, DB1998, and A 4
Normalized Absorbance
Temperature (°C)
DNA
DB224 (2:1)
DB323 (2:1)
DB1255 (2:1)
DB1998 (2:1)

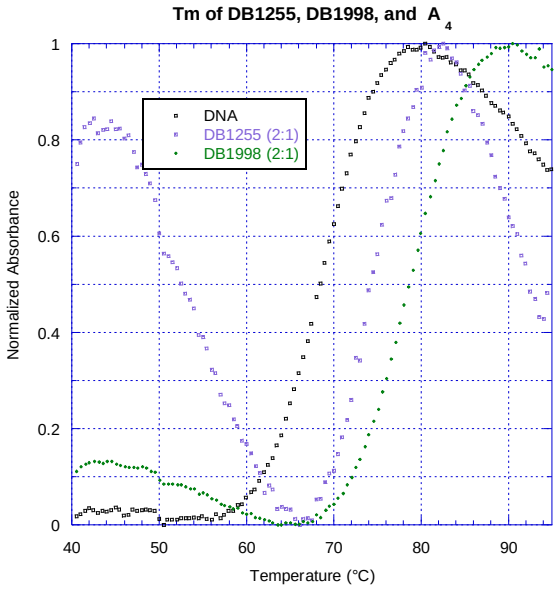

Tm of DB232 and A 4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

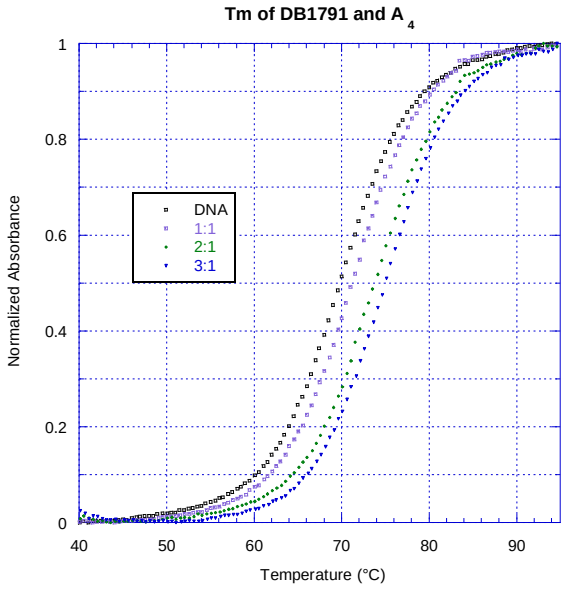

Tm of DB321 and A 4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

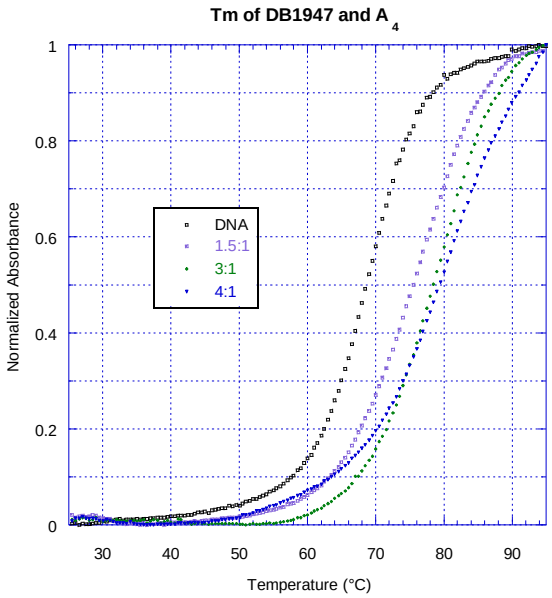

Tm of DB325, DB496, DB2247, DB2260, and A 4
Normalized Absorbance
Temperature (°C)
DNA
DB325 (2:1)
DB496 (2:1)
DB2247 (2:1)
DB2260 (2:1)

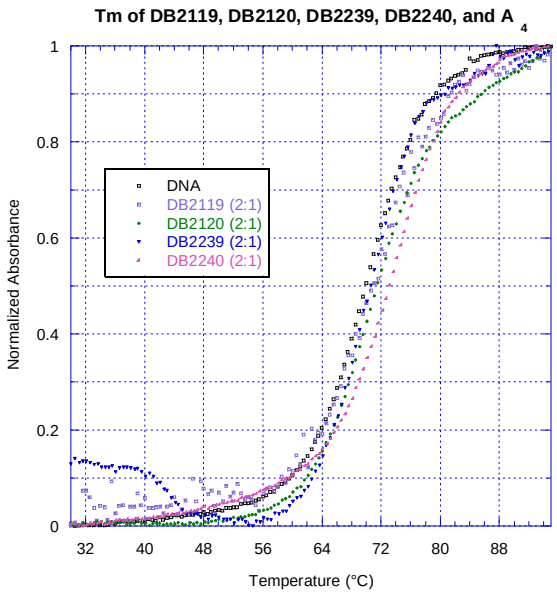

Tm of DB462 and A4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

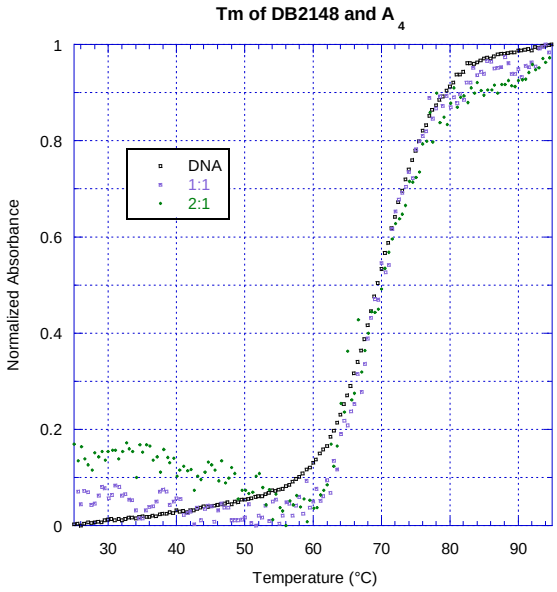

Tm of DB510 and A4
Normalized Absrobance
Temperature (°C)
DNA
1:1
2:1
3:1

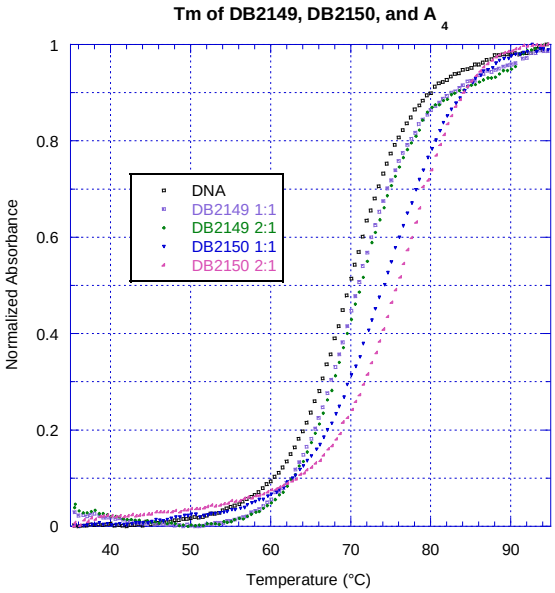

Tm of DB1255, DB1998, and A4
Normalized Absorbance
Temperature (°C)
DNA
DB1255 (2:1)
DB1998 (2:1)

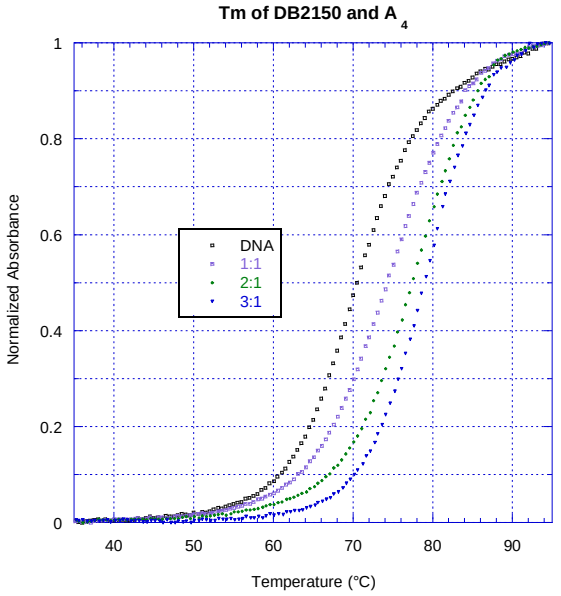

Tm of DB1791 and A4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1

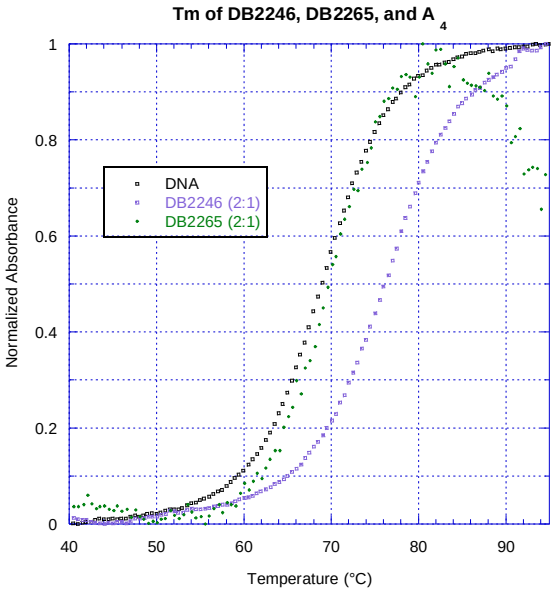

Tm of DB1947 and A 4
Normalized Absorbance
Temperature (°C)
DNA
1.5:1
3:1
4:1

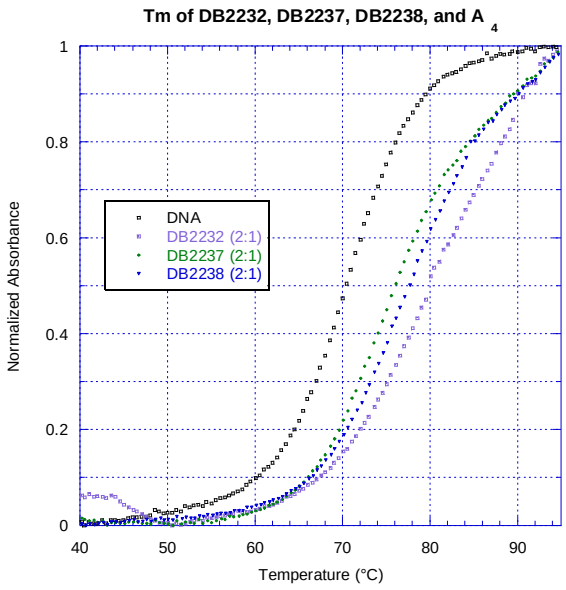

Tm of DB2119, DB2120, DB2239, DB2240, and A 4
Normalized Absorbance
Temperature (°C)
DNA
DB2119 (2:1)
DB2120 (2:1)
DB2239 (2:1)
DB2240 (2:1)


Tm of DB2148 and A 4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1


Tm of DB2149, DB2150, and A 4
Normalized Absorbance
Temperature (°C)
DNA
DB2149 1:1
DB2149 2:1
DB2150 1:1
DB2150 2:1


Tm of DB2150 and A4
Normalized Absorbance
Temperature (°C)
DNA
1:1
2:1
3:1


Tm of DB2246, DB2265, and A4
Normalized Absorbance
Temperature (°C)
DNA
DB2246 (2:1)
DB2265 (2:1)


Tm of DB2232, DB2237, DB2238, and A4
Normalized Absorbance
Temperature (°C)
DNA
DB2232 (2:1)
DB2237 (2:1)
DB2238 (2:1)